

(Translation)

PATENT OFFICE
JAPANESE GOVERNMENT



This is to certify that the annexed is a true copy of
the following application as filed with this Office.

Date of Application: February 23, 2001

Application Number: Japanese Patent Application
No. 047762/2001

Applicant(s): RIKEN

July 27, 2001

Commissioner,
Patent Office

Kozo Oikawa (seal)

Certificate No. 2001-3066119

#4

日 本 国 特 許 庁
JAPAN PATENT OFFICE

JC971 U.S. PTO
09/938901
08/24/01

別紙添付の書類に記載されている事項は下記の出願書類に記載されている事項と同一であることを証明する。

This is to certify that the annexed is a true copy of the following application as filed with this Office

出 願 年 月 日

Date of Application:

2001年 2月23日

出 願 番 号

Application Number:

特願2001-047762

出 願 人

Applicant(s):

理化学研究所

2001年 7月27日

特 許 庁 長 官
Commissioner,
Japan Patent Office

及 川 耕 造

出証番号 出証特2001-3066119

【書類名】 特許願

【整理番号】 RJH12-082T

【特記事項】 特許法第 3 0 条第 1 項の規定の適用を受けようとする特
許出願

【提出日】 平成13年 2月23日

【あて先】 特許庁長官 殿

【国際特許分類】 C12N 15/00

【発明の名称】 D N A 修復酵素遺伝子

【請求項の数】 11

【発明者】

 【住所又は居所】 兵庫県佐用郡三日月町光都 1 丁目 1 番 1 号 理化学研究
 所 播磨研究所内

 【氏名】 倉光 成紀

【発明者】

 【住所又は居所】 神奈川県横浜市鶴見区末広町 1 丁目 7 番 2 2 号 理化学
 研究所 横浜研究所内

 【氏名】 横山 茂之

【特許出願人】

 【識別番号】 000006792

 【氏名又は名称】 理化学研究所

【代理人】

 【識別番号】 100091096

 【弁理士】

 【氏名又は名称】 平木 祐輔

【選任した代理人】

 【識別番号】 100096183

 【弁理士】

 【氏名又は名称】 石井 貞次

BEST AVAILABLE COPY

【手数料の表示】

【予納台帳番号】 015244

【納付金額】 21,000円

【提出物件の目録】

【物件名】 明細書 1

【物件名】 図面 1

【物件名】 要約書 1

【包括委任状番号】 9503608

【プルーフの要否】 要

【書類名】 明細書

【発明の名称】 DNA修復酵素遺伝子

【特許請求の範囲】

【請求項1】 以下の(a)又は(b)のタンパク質。

(a) 配列番号2、4、6若しくは8に示されるアミノ酸配列を含むタンパク質

(b) 配列番号2、4、6若しくは8に示されるアミノ酸配列において1若しくは数個のアミノ酸が欠失、置換若しくは付加されたアミノ酸配列を含み、かつ、DNA修復酵素活性を有するタンパク質

【請求項2】 以下の(a)又は(b)のタンパク質をコードするDNAを含むDNA修復酵素遺伝子。

(a) 配列番号2、4、6若しくは8に示されるアミノ酸配列を含むタンパク質

(b) 配列番号2、4、6若しくは8に示されるアミノ酸配列において1若しくは数個のアミノ酸が欠失、置換若しくは付加されたアミノ酸配列を含み、かつ、DNA修復酵素活性を有するタンパク質

【請求項3】 以下の(c)又は(d)のDNAを含むDNA修復酵素遺伝子。

(c) 配列番号1、3、5若しくは7に示される塩基配列を含むDNA

(d) 配列番号1、3、5若しくは7に示される塩基配列を含むDNAとストリンジントな条件下でハイブリダイズし、かつ、DNA修復酵素活性を有するタンパク質をコードするDNA

【請求項4】 請求項2又は3記載のDNA修復酵素遺伝子を含む組換えベクター。

【請求項5】 請求項4記載の組換えベクターを含む形質転換体。

【請求項6】 請求項5記載の形質転換体を培養し、得られる培養物からDNA修復酵素を採取することを特徴とするDNA修復酵素の製造方法。

【請求項7】 請求項1記載のタンパク質の存在下でDNA合成反応を行うことを特徴とする、DNAのエラー配列を修復する方法。

【請求項8】 請求項1記載のタンパク質の存在下でDNA合成反応を行うことを特徴とする、DNA配列のエラー合成を防止する方法。

【請求項9】 請求項2又は3記載の遺伝子中に修飾遺伝子が組み込まれた

遺伝子を宿主に導入してなる修復遺伝子破壊株。

【請求項 1 0】 修飾遺伝子がマーカー遺伝子である請求項 9 記載の修復遺伝子破壊株。

【請求項 1 1】 宿主が好熱菌である請求項 9 又は 10 記載の修復遺伝子破壊株。

【発明の詳細な説明】

【 0 0 0 1 】

【発明の属する技術分野】

本発明は、DNA修復酵素、DNA修復遺伝子及びDNA修復方法に関する。

【 0 0 0 2 】

【従来 of 技術】

生命の維持に必要な全ての情報が書き込まれている細胞内のゲノムDNAは、内外様々な要因により、常に損傷を受けている。外的な要因としては紫外線、電離放射線、環境中に存在する化学物質などがあり、内的なものとしては、エネルギー代謝や酸化ストレスから生ずる活性酸素などが挙げられる。また、DNAの複製が鋳型どおりに行われずにミスマッチが生じる場合がある。

【 0 0 0 3 】

これらの損傷部やミスマッチが放置されると、この部分の塩基が本来のものと異なったものに変化することになり、遺伝情報の狂い、すなわち突然変異が生じる。タンパク質のコード領域中に突然変異が生じた場合、本来のものよりも活性が低いか、あるいは活性を持たないタンパク質が作られることがある。また、このタンパク質が全く生産されなくなることもある。突然変異が遺伝子調節領域に生じると、この領域に制御されるタンパク質の合成量が、異常に増加又は減少する。さらに、他のタンパク質による制御が効かなくなる場合もある。これらの変化は、その細胞にアポトーシスを引き起こしたり、その細胞を癌化させたりする原因となる。

【 0 0 0 4 】

このようにDNAの損傷やミスマッチは、その細胞自体の、さらにはその細胞が属する個体の生死に関わるものであるため、細胞はこれを修復し、遺伝情報を正

確に維持するためのメカニズムを有している。これがDNA修復機構と呼ばれるものである。DNA修復機構には、塩基除去修復、光回復、ヌクレオチド除去修復、ミスマッチ修復、組換え修復などの種類がある。DNA修復機構を解明することにより、癌を始めとする疾病の研究、環境因子が生体に与える影響の研究などに有用な知見が得られるものと期待される。また、修復機構に関わるある種のタンパク質については、今や分子生物学の分野を越えた様々な分野で重要な技術となっているPCRの精度をより高くする効果も期待される。

【 0 0 0 5 】

既に幾つかのDNA修復酵素が様々な生物からクローニングされ、その中のあるものについてはタンパク質の立体構造解析も行われている。しかし、多くのものは、現在のところ、遺伝学的研究が主であり、生化学的研究はほとんど行われていない。今後、修復機構を明らかにし、医学を始め様々な分野で有用な知見を得るためには、DNA修復に関わるすべての遺伝子をクローン化し、タンパク質の立体構造解析、さらには詳細な機能解析を進める必要がある。

【 0 0 0 6 】

【発明が解決しようとする課題】

本発明は、DNA修復酵素、DNA修復遺伝子及びDNA修復方法を提供することを目的とする。

【 0 0 0 7 】

【課題を解決するための手段】

本発明者は、上記課題を解決するため鋭意研究を行った結果、高度好熱菌からDNA修復酵素遺伝子を単離することに成功し、本発明を完成するに至った。

すなわち、本発明は以下の通りである。

【 0 0 0 8 】

(1) 以下の(a)又は(b)のタンパク質。

(a) 配列番号 2、4、6 若しくは 8 に示されるアミノ酸配列を含むタンパク質

(b) 配列番号 2、4、6 若しくは 8 に示されるアミノ酸配列において1若しくは数個のアミノ酸が欠失、置換若しくは付加されたアミノ酸配列を含み、かつ、DNA修復酵素活性を有するタンパク質

【 0 0 0 9 】

(2) 以下の(a)又は(b)のタンパク質をコードするDNAを含むDNA修復酵素遺伝子。

(a) 配列番号 2、4、6 若しくは 8 に示されるアミノ酸配列を含むタンパク質

(b) 配列番号 2、4、6 若しくは 8 に示されるアミノ酸配列において1若しくは数個のアミノ酸が欠失、置換若しくは付加されたアミノ酸配列を含み、かつ、DNA修復酵素活性を有するタンパク質

【 0 0 1 0 】

(3) 以下の(c)又は(d)のDNAを含むDNA修復酵素遺伝子。

(c) 配列番号 1、3、5 若しくは 7 に示される塩基配列を含むDNA

(d) 配列番号 1、3、5 若しくは 7 に示される塩基配列を含むDNAとストリンジントな条件下でハイブリダイズし、かつ、DNA修復酵素活性を有するタンパク質をコードするDNA

(4) 前記DNA修復酵素遺伝子を含む組換えベクター。

(5) 前記組換えベクターを含む形質転換体。

(6) 前記形質転換体を培養し、得られる培養物からDNA修復酵素を採取すること
を特徴とするDNA修復酵素の製造方法。

【 0 0 1 1 】

(7) 前記タンパク質の存在下でDNA合成反応を行うことを特徴とする、DNAのエラー配列を修復する方法。

(8) 前記タンパク質の存在下でDNA合成反応を行うことを特徴とする、DNA配列のエラー合成を防止する方法。

(9) 前記DNA修復酵素遺伝子中に修飾遺伝子が組み込まれた遺伝子を宿主に導入してなる修復遺伝子破壊株。

【 0 0 1 2 】

修飾遺伝子としてはマーカー遺伝子が挙げられ、宿主としては好熱菌が挙げられる。また、前記タンパク質は4℃～100℃、好ましくは80℃まで、さらに好ましくは75℃まで安定性を有するものである。

以下、本発明を詳細に説明する。

【 0 0 1 3 】

【発明の実施の形態】

前述の通り、遺伝子の修復機構を明らかにして様々な分野で有用な知見を得るためには、優れた安定性を有する高度好熱菌由来のDNA修復酵素遺伝子を数多くクローニングすることが重要である。本発明は、熱安定性が高く、立体構造解析又は分子機能解析に適したサーマス属に属する高度好熱菌、特にサーマス・サーモフィラス (*Thermus thermophilus*) 由来のDNA修復酵素遺伝子を用いて、DNA修復酵素タンパク質の量産化及び基質認識機構の解析等を行い完成させたものである。

【0014】

1. DNA修復酵素

自然界においては、エネルギー代謝や酸化ストレスから生ずる活性酸素等の内因、あるいは紫外線、電離放射線、化学物質等の種々の外因が存在し、これらによってDNAは損傷を受けている。また、DNA複製は鋳型通りに行われず塩基間のミスマッチが生じる場合があり、例えばポリメラーゼ連鎖反応に使用するDNAポリメラーゼの種類によっては鋳型通りにDNAを合成されないことがある。本発明のタンパク質は、これらのミスマッチを修復し、適正な塩基対合となるように修復する酵素である。

本発明において単離されたDNA修復酵素は以下の4種類 (MutY、RecJ、RecF及びTRCF) である。

【0015】

(1) MutY

好気性生物のDNAは、エネルギー代謝やストレスから生ずる活性酸素で絶えず損傷を受けている。グアニンは、8-オキシグアニン (GO) へと酸化されやすく、シトシンばかりでなくアデニンとも塩基対を作り、C:G→A:Tトランスバージョンを引き起こす (図1)。この突然変異を防ぐために、MutYはA:GOのミスマッチを認識し、アデニンを除去する。また、G:GOのミスマッチを認識してグアニンを除去する。

【0016】

作用：図2A～Eに示す。まず、傷害を含むDNAの損傷部位から、DNAグリコシラ

ーゼ活性により塩基部分を除去する (A)。次に、APリアーゼ活性により、塩基部分が除去された部位 (AP部位) の3' 側を切断する (B)。最終的にエステラーゼ、DNAポリメラーゼ、DNAリガーゼの作用によりギャップが埋められて修復が完了する (E)。

【 0 0 1 7 】

分子量：アミノ酸配列から計算した理論的分子量は36kDaであり、SDS-ポリアクリルアミドゲル電気泳動により推定される分子量は約36kDaであり (図3)、ゲルろ過 (Superdex 200HR, 50mM Tris-HCl (pH8), 0.5M NaCl) により見積もられる分子量は31kDaである (図4)。

アミノ酸配列：配列番号2に示される。他の微生物由来のMutYのアミノ酸配列を比較すると、N-グリコシラーゼ活性に必須な部分及び鉄硫黄クラスターを構成する部分が保存されていることが分かる (図5)。

【 0 0 1 8 】

基質特異性：AとGとのミスマッチ (A:G)、AとGとのミスマッチ (A:G)、GとGとのミスマッチ (G:G) を認識し、除去する (図7)。

吸収スペクトル：50mM リン酸カリウム pH7.5、0.8M KCl、1mM DTT、1mM EDTA 及び10%グリセロールを含む溶液中において測定した結果、410nm付近に鉄硫黄クラスターに特有なスペクトルを有する (図8)。

【 0 0 1 9 】

α -ヘリックス含有率：50mM Tris-HCl pH8.0、0.1M KCl、1mM DTE、1mM EDTA 及び20%グリセロールを含む溶液中におけるCDスペクトル解析によると、約40%の α ヘリックスを有する (図9)。

熱安定性：50mM リン酸カリウム pH7.5、0.1M KCl、1mM DTE、1mM EDTA 及び20%グリセロールを含む溶液中で、測定温度を変えてCDスペクトルを解析したところ、中性条件 (pH7.5) で24℃～80℃で (特に75℃まで) 安定である (図10)。

【 0 0 2 0 】

(2) RecJ

RecJは、一本鎖DNAに特異的なエキソヌクレアーゼ活性、及びデオキシリボジエステラーゼ活性の両方の活性を有し、塩基除去修復及びミスマッチ修復の両方

の修復系に関わるDNA修復酵素である（図12）。また、RecQ及びSSB（いずれも一本鎖DNA結合タンパク質）と協調して相同組換えの初期過程を担うことも知られている。

【0021】

塩基除去修復をする場合の機能は、DNAグリコシラーゼ及びAPエンドヌクレアーゼの作用により生じたニック部分の3'側を切断する（図12左）。

ミスマッチ修復をする場合の機能は、MutS、MutH、MutL又はUvrDの作用により生じた一本鎖DNA部分を5'→3'方向へ分解する（図12中央）。

【0022】

相同組換えに働く場合の機能は、RecQ又はSSBの作用により生じた一本鎖DNA部分を5'→3'方向へ分解する（図12右）。

RecJのホモログは、原核生物のみならず、酵母、ショウジョウバエなどの真核生物にまで広く存在し、特徴的なモチーフを共通して有する。そのモチーフを有するタンパク質は、原核生物のみならず、酵母やショウジョウバエにも存在する（図14）。

【0023】

作用：RecJは一本鎖DNAを5'→3'方向にのみ分解するエキソヌクレアーゼ活性を有する（図12）。

分子量：アミノ酸配列から計算した理論的分子量は約50kDaであり、SDS-ポリアクリルアミドゲル電気泳動により推定される分子量は約50kDaである（図13）。

アミノ酸配列：配列番号4に示される。

【0024】

基質特異性：一本鎖DNAに対して特異性を有し、 $K_m=6.2\mu M$ である（図17～23）。

α -ヘリックス含有率：50mM K-Pi、100mM KCl、0.1mM DTE、及び0.1mM EDTA(pH7.2)を含む溶液中におけるCDスペクトル解析によると、約50%の α ヘリックスを有する（図15）。

熱安定性：1.6 μM RecJ、50mM K-Pi、100mM KCl、0.1mM DTE及び0.1mM EDTA p

H7.2を含む溶液中で、 $1.6\mu\text{M}$ RecJを用い、測定温度を変化させてCDスペクトルを測定した結果、 60°C まで安定である（図16）。

【 0 0 2 5 】

(3) RecF

RecFタンパク質は、これまでの遺伝学的解析によって、DNA組換え修復、遺伝的組換え、さらにDNA複製に重要な機能を果たすことが知られている。

作用：RecO及びRecRタンパク質と協同して、傷害部位での複製を防ぐ（図24）。すなわち、DNAに傷害が入り(A)、傷害部位で複製複合体の反応が停止すると(B)、RecF、RecO及びRecRタンパク質の複合体がDNAに結合する(C)。その後新たな複製が開始され(D)、RecAが相同領域の対合を起こし(E)、鎖交換反応の進行とDNA合成が行われる(F)。そして、相同鎖間の対合によりできたホリデイ(Holiday)構造（二本鎖DNAのそれぞれの鎖に相同な娘鎖が互いに対合している構造）をRuvA、RuvB及びRuvCが解消して修復を完了させる(G)。

【 0 0 2 6 】

分子量：アミノ酸配列から計算した理論的分子量は 37.8kDa であり、ゲルろ過(Superdex 200HR, 50mM Tris-HCl, 2.0M KCl(pH7.5))により見積もられる分子量は 22kDa であり（図26）、SDS-PAGEにより見積もられる分子量は 37kDa である（図25）。

アミノ酸配列：配列番号8に示される。他の微生物由来のRecFのアミノ酸配列を比較すると、部分的に高いホモロジーを有することがわかる（図27）。

【 0 0 2 7 】

基質特異性： K_m 値は、 37°C で $31\mu\text{M}$ 、 25°C で $32\mu\text{M}$ である。

α -ヘリックス含有率： 50mM Tris-HCl、 100mM KCl (pH7.2)を含む溶液中におけるCDスペクトル解析によると、約40%の α ヘリックスを有する。

熱安定性：CDスペクトルを解析した結果、pH7.5で約 50°C まで安定である。

ATPase活性：RecF単独でもATPase活性を有し（図29）、一本鎖DNAでは活性が増加し、二本鎖DNAでは活性が減少する（図30）。

等電点： 9.1

【 0 0 2 8 】

(4) TRCF

UvrA、UvrB、UvrCによって担われるヌクレオチド除去修復は、最も広い範囲でDNA傷害を認識・除去することができる機構である。このうちUvrAとUvrBはUvrAB複合体を形成し、DNA傷害を特異的に認識する。UvrBの立体構造解析の結果、UvrAと相互作用すると考えられている領域が β シートからなる1つのドメイン (UvrB- β) を形成していることが分かっている (図32) (Nakagawa et al., J. Biochem. 126, 986-990, 1999)。TRCF (transcription-repair coupling factor) は、UvrAと相互作用し、傷害を含む転写鎖のDNA修復を促進する因子である。TRCFには、UvrB- β のアミノ酸配列と相同な領域 (TRCF- β) が存在し、この領域がUvrA結合部位であると考えられている。

【0029】

作用：UvrAと相互作用し、傷害を含む転写鎖のDNA修復を促進する (図31)。原核生物のヌクレオチド除去修復機構は以下の通りである (図31)。UvrAとUvrBとの複合体が傷害部位を認識し結合する。また、転写鎖の傷害はTRCF及びUvrAによって認識される。さらに、UvrCの作用により傷害部位の両端が切断され除去される。この後、UvrD (ヘリカーゼII)、DNAポリメラーゼI及びDNAリガーゼの作用により修復合成が完了する。

【0030】

分子量：アミノ酸配列から計算した理論的分子量は37.8kDaであり、またUvrAとの結合部位であると考えられるTRCF- β 領域の分子量は14.4kDaである。SDS-ポリアクリルアミドゲル電気泳動により推定されるTRCF- β 領域の分子量は14.4kDaである (図33、下パネル)。

【0031】

アミノ酸配列：配列番号6に示される。UvrBとTRCFとの相同領域 (UvrB- β , TRCF- β) のアミノ酸配列は、高度に保存されている (図34)。

CDスペクトル：50mM Tris-HCl、100mM KCl (pH7.9) 中において測定したTRCF- β のCDスペクトルは、同じ条件で測定したUvrB- β のCDスペクトルと類似している (図35)。

【0032】

熱安定性：50mM Tris-HCl、100mM KCl(pH7.9)中においてTRCF- β のCDスペクトルを測定した結果、20℃～75℃で安定である（図36）。

pH安定性：100mM KClを含む各種緩衝液中においてCDスペクトルを測定した結果、25℃でpH4～9の範囲で安定である（図37）。

BIACoreによるUvrAとの相互作用解析の結果、TRCFはUvrAと結合することが示され、この結合の解離定数は、ATP存在下では0.5 μ M、ATP非存在下では1.3 μ Mである（図38、39）。

【 0 0 3 3 】

2. 遺伝子のクローニング

本発明の遺伝子は、前記DNA修復酵素をコードするものであり、以下のクローニング手法により得ることができる。本発明の遺伝子のクローニングについて具体的に説明する。

本発明の遺伝子は、高度好熱菌であるサーマス・サーモフィラス(*Thermus thermophilus*)のゲノムDNAから単離することができる。

【 0 0 3 4 】

(1) ゲノムDNAの調製

ゲノムDNAは、上記菌体から通常行われる手法に従って調製することができる。例えば、グアニジンを含むバッファーで菌体を破碎し、フェノール抽出により粗DNA分画を得た後、塩化セシウム密度勾配超遠心法により、精製されたゲノムDNAを得る。上記の手法により得られたゲノムDNAを適当な制限酵素（例えばEcoRI、BamHI、Sau3AI等）で切断する。DNA断片を連結させるには、例えばT4DNAリガーゼを用いる。

【 0 0 3 5 】

上記制限酵素処理されたDNA断片は、当該処理に使用した制限酵素と同種類の酵素（例えばEcoRI、BamHI等）、あるいは当該処理に使用した制限酵素の切断部位に相補的な接着末端を生じる制限酵素（例えばSau3AIに対してBamHI）で切断したベクターとライゲーションを行う。これを用いてライブラリーを作製することもできる。上記DNA断片はPCR等により増幅しておくこともできる。ベクターには、宿主微生物で自律的に増殖し得るファージ又はプラスミドが使用される。フ

ファージベクターとしては、例えばEMBL3、M13、 λ gt11等が挙げられ、プラスミドベクターとしては、例えばpET系（pET-3a等）、pBR系（pBR322等）、pUC系（pUC18等）、pBluescript II（STRATAGENE社製）等が挙げられる。さらに、大腸菌や枯草菌などの2種以上の宿主微生物で自律的増殖が可能なベクターのほか、各種のシャトルベクターを使用することもできる。DNA断片とベクター断片とを連結させるには、公知のDNAリガーゼ（例えばT4DNAリガーゼ）を用いる。そして、DNA断片とベクター断片とをアニーリングさせた後連結させ、これを宿主微生物に導入する。宿主微生物へのDNAの導入は、公知の任意の方法を採用することができる。例えば、宿主微生物が大腸菌の場合は、電気穿孔法、カルシウム法等を採用することができる。ファージDNAを大腸菌に導入する場合はキットを用いたインビトロ・パッケージング法（Gigapack II；STRATAGENE社製）等を採用することができる。

【 0 0 3 6 】

次に、抗生物質含有培地で宿主をスクリーニングし、選抜された宿主からアルカリ-SDS法等によってプラスミドを回収することにより、本発明の遺伝子を含むゲノムDNA断片が得られる。

得られたDNAの塩基配列の決定方法は特に限定されるものではない。例えば、Perkin Elmer社製の蛍光ダイデオキシターミネーターを含有するPRISMシーケンシングキット等を使って反応を行い、Applied Biosystem社製のオートシーケンサー（例えばモデルABI377）等で塩基配列を決定する。

【 0 0 3 7 】

本発明において、修復酵素遺伝子としてMutY、RecJ、RecF及びTRCFが得られた。配列番号1にMutY遺伝子の塩基配列を、配列番号2にMutY遺伝子によりコードされるアミノ酸配列を例示する。配列番号3にRecJ遺伝子の塩基配列を、配列番号4にRecJ遺伝子によりコードされるアミノ酸配列を例示する。配列番号5にRecF遺伝子の塩基配列を、配列番号6にRecF遺伝子によりコードされるアミノ酸配列を例示する。そして、配列番号7にTRCF遺伝子の塩基配列を、配列番号8にTRCF遺伝子によりコードされるアミノ酸配列を例示する。但し上記各アミノ酸配列を含むタンパク質がDNA修復酵素活性を有する限り、また、4℃～100℃、好まし

くは80℃まで、さらに好ましくは75℃まで安定性を有する限り、当該アミノ酸配列において複数個、好ましくは1若しくは数個のアミノ酸に欠失、置換、付加等の変異が生じてもよい。

【0038】

例えば、配列番号2、4、6又は8に示されるアミノ酸配列の1～10個、好ましくは1～5個のアミノ酸が欠失してもよく、配列番号2、4、6又は8に示されるアミノ酸配列に1～10個、好ましくは1～5個のアミノ酸が付加してもよく、あるいは、配列番号2、4、6又は8に示されるアミノ酸配列の1～10個、好ましくは1～5個のアミノ酸が他のアミノ酸に置換してもよい。このように、変異体を作製することによって熱に対してより安定なタンパク質を得ることが可能である。

【0039】

さらに、配列番号2に示されるアミノ酸配列と少なくとも60%、好ましくは少なくとも80%、さらに好ましくは少なくとも90%のホモロジーを有し、かつDNA修復酵素活性を有するタンパク質も、本発明のタンパク質に含まれる。

ここで、DNA修復酵素活性とは、DNAに生じた各種傷害、及びこの傷害から生じたミスマッチ部位を認識、除去し、ギャップを埋めることができる活性を意味し、修復の対象となる傷害として、活性酸素による傷害、紫外線照射により生ずる傷害、化学物質により生ずる傷害、PCRエラー等が挙げられる。

【0040】

また、「安定性」とは、4℃～100℃の範囲で、好ましくは80℃まで、さらに好ましくは75℃まで、CDスペクトル等で測定したタンパク質の構造が変化しないことを意味する。

また、DNA修復酵素遺伝子とストリンジントな条件下でハイブリダイズすることができるDNAも本発明の遺伝子に含まれる。ストリンジントな条件とは、ナトリウム濃度が15～300mM、好ましくは15～75mMであり、温度が50～60℃、好ましくは55～60℃での条件をいう。

【0041】

一旦本発明の遺伝子の塩基配列が確定されると、その後は化学合成によって、

又はクローニングされたcDNAを鋳型としたPCRによって、あるいは該塩基配列を有するDNA断片をプローブとしてハイブリダイズさせることによって、本発明の遺伝子を得ることができる。さらに、部位特異的突然変異誘発法等によって本発明の遺伝子の変異型であってDNA修復酵素活性を有するタンパク質を発現し得る遺伝子を合成することもできる。

【 0 0 4 2 】

なお、遺伝子に変異を導入するには、Kunkel法、Gapped duplex法等の公知の手法又はこれに準ずる方法を採用することができる。例えば部位特異的突然変異誘発法を利用した変異導入用キット（例えばMutan-K(TAKARA社製)やMutan-G(TAKARA社製)）などを用いて、あるいは、TAKARA社のLA PCR in vitro Mutagenesis シリーズキットを用いて変異の導入が行われる。

【 0 0 4 3 】

3. 組換えベクター及び形質転換体の作製

(1) 組換えベクターの作製

本発明の組換えベクターは、適当なベクターに本発明の遺伝子を連結(挿入)することにより得ることができる。本発明の遺伝子を挿入するためのベクターは、宿主中で複製可能なものであれば特に限定されず、例えば、プラスミド DNA、ファージ DNA等が挙げられる。

【 0 0 4 4 】

プラスミド DNAとしては、大腸菌由来のプラスミド（例えばpBR322、pBR325、pUC118、pUC119、pUC18、pUC19、pBluescript等）、枯草菌由来のプラスミド（例えばpUB110、pTP5等）、酵母由来のプラスミド（例えばYEp系プラスミド、YCp系プラスミド等）などが挙げられ、ファージDNAとしてはλファージ（Charon4A、Charon21A、EMBL3、EMBL4、λgt10、λgt11、λZAP等）が挙げられる。さらに、レトロウイルス又はワクシニアウイルスなどの動物ウイルス、バキュロウイルスなどの昆虫ウイルスベクターを用いることもできる。

【 0 0 4 5 】

ベクターに本発明の遺伝子を挿入するには、まず、精製されたDNAを適当な制限酵素で切断し、適当なベクター DNAの制限酵素部位又はマルチクローニングサ

イトに挿入してベクターに連結する方法などが採用される。

本発明の遺伝子は、その遺伝子の機能が発揮されるようにベクターに組み込まれることが必要である。そこで、本発明のベクターには、プロモーター、本発明の遺伝子のほか、所望によりエンハンサーなどのシスエレメント、スプライシングシグナル、ポリA付加シグナル、選択マーカー、リボソーム結合配列（SD配列）などを連結することができる。なお、選択マーカーとしては、例えばジヒドロ葉酸還元酵素遺伝子、アンピシリン耐性遺伝子、ネオマイシン耐性遺伝子等が挙げられる。

【 0 0 4 6 】

(2) 形質転換体の作製

本発明の形質転換体は、本発明の組換えベクターを、目的遺伝子が発現し得るように宿主中に導入することにより得ることができる。ここで、宿主としては、本発明の遺伝子を発現できるものであれば特に限定されるものではなく、原核生物、真核生物、植物、動物細胞、昆虫細胞等が挙げられる。

【 0 0 4 7 】

原核生物としては、大腸菌(*Escherichia coli*)等のエッシェリヒア属、バチルス・ズブチリス(*Bacillus subtilis*)等のバチルス属、シュードモナス・プチダ(*Pseudomonas putida*)等のシュードモナス属に属する細菌などを例示することができる。

真核生物としては、サッカロミセス・セレビシエ(*Saccharomyces cerevisiae*)、シゾサッカロミセス・ポンベ(*Schizosaccharomyces pombe*)、ピキア・パストリス(*Pichia pastris*)等の酵母を例示することができる。

【 0 0 4 8 】

動物細胞としては、サルのCOS細胞、チャイニーズハムスター卵巣細胞（CHO細胞）、Vero細胞、マウスL細胞等を例示することができる。

昆虫細胞としては、Sf9等を例示することができる。

植物細胞としては、植物体全体、植物器官(例えば葉、花卉、茎、根、種子等)、植物組織(例えば表皮、師部、柔組織、木部、維管束等)又は植物培養細胞（プロトプラストを含む）を例示することができる。

【 0 0 4 9 】

大腸菌、酵母等の細菌を宿主とする場合は、本発明の組換えベクターが該細菌中で自律複製可能であると同時に、本発明の遺伝子のほか、プロモーター、リボソーム結合配列、転写終結配列、プロモーターを制御する遺伝子等が含まれていてもよい。

【 0 0 5 0 】

プロモーターは、大腸菌等の宿主中で発現できるものであればいずれを用いてもよい。例えばtrpプロモーター、lacプロモーター、 P_L プロモーター、 P_R プロモーターなどの、大腸菌やファージに由来するプロモーターが用いられる。tacプロモーターなどのように、人為的に設計改変されたプロモーターを用いてもよい。細菌への組換えベクターの導入方法は、細菌にDNAを導入する方法であれば特に限定されるものではない。例えばカルシウムイオンを用いる方法、エレクトロポレーション法等が挙げられる。

【 0 0 5 1 】

酵母を宿主とする場合は、プロモーターは酵母中で発現できるものであれば特に限定されず、例えばGAL1プロモーター、GAL10プロモーター、ヒートショックタンパク質プロモーター、MF α 1プロモーター、PH05プロモーター、PGKプロモーター、GAPプロモーター、ADHプロモーター、AOX1プロモーター等を用いることができる。酵母への組換えベクターの導入方法は、酵母にDNAを導入する方法であれば特に限定されず、例えばエレクトロポレーション法、スフェロプラスト法、酢酸リチウム法等が挙げられる。

【 0 0 5 2 】

動物細胞を宿主とする場合は、プロモーターとしてSR α プロモーター、SV40プロモーター、LTRプロモーター、CMVプロモーター等が用いられ、また、ヒトサイトメガロウイルスの初期遺伝子プロモーター等を用いてもよい。動物細胞への組換えベクターの導入方法としては、例えばエレクトロポレーション法、リン酸カルシウム法、リポフェクション法等が挙げられる。

【 0 0 5 3 】

昆虫細胞を宿主とする場合は、昆虫細胞への組換えベクターの導入方法として

は、例えばリン酸カルシウム法、リポフェクション法、エレクトロポレーション法などが挙げられる。

宿主が植物であるときの形質転換に用いられる植物としては、アブラナ科、イネ科、ナス科、マメ科等に属する植物（下記参照）が挙げられるが、これらの植物に限定されるものではない。

【 0 0 5 4 】

アブラナ科：シロイヌナズナ (*Arabidopsis thaliana*)

ナス科：タバコ (*Nicotiana tabacum*)

イネ科：トウモロコシ (*Zea mays*)、イネ (*Oryza sativa*)

マメ科：ダイズ (*Glycine max*)

【 0 0 5 5 】

上記組換えベクターは、通常の形質転換方法、例えば電気穿孔法（エレクトロポレーション法）、アグロバクテリウム法、パーティクルガン法、ポリエチレングリコール（PEG）法等によって植物中に導入することができる。

形質転換の結果得られる腫瘍組織やシュート、毛状根などは、そのまま細胞培養、組織培養又は器官培養に用いることが可能であり、また従来知られている植物組織培養法を用い、適当な濃度の植物ホルモン（オーキシン、サイトカイニン、ジベレリン、アブシジン酸、エチレン、ブラシノライド等）の投与などにより植物体に再生させることができる。

【 0 0 5 6 】

遺伝子が宿主に組み込まれたか否かの確認は、PCR法、サザンハイブリダイゼーション法、ノーザンハイブリダイゼーション法等により行うことができる。例えば、形質転換体からDNAを調製し、DNA特異的プライマーを設計してPCRを行う。PCRは、前記プラスミドを調製するために使用した条件と同様の条件で行われる。その後は、増幅産物についてアガロースゲル電気泳動、ポリアクリルアミドゲル電気泳動又はキャピラリー電気泳動等を行い、臭化エチジウム、SYBR Green 液等により染色し、そして増幅産物をバンドとして検出することにより、形質転換されたことを確認する。また、予め蛍光色素等により標識したプライマーを用いてPCRを行い、増幅産物を検出することもできる。さらに、マイクロプレート

等の固相に増幅産物を結合させ、蛍光又は酵素反応等により増幅産物を確認する方法も採用してもよい。

【0057】

4. DNA修復酵素タンパク質の生産

本発明のタンパク質は、前記形質転換体を培養し、その培養物から採取することにより得ることができる。「培養物」とは、培養上清のほか、培養細胞若しくは培養菌体又は細胞若しくは菌体の破砕物のいずれをも意味するものである。

本発明の形質転換体を培養する方法は、宿主の培養に用いられる通常の方法に従って行われる。

【0058】

大腸菌や酵母菌等の微生物を宿主として得られた形質転換体を培養する培地は、微生物が資化し得る炭素源、窒素源、無機塩類等を含有し、形質転換体の培養を効率的に行うことができる培地であれば、天然培地、合成培地のいずれを用いてもよい。炭素源としては、グルコース、フラクトース、スクロース、デンプン等の炭水化物、酢酸、プロピオン酸等の有機酸、エタノール、プロパノール等のアルコール類が挙げられる。窒素源としては、アンモニアのほか、塩化アンモニウム、硫酸アンモニウム、酢酸アンモニウム、リン酸アンモニウム等の無機酸若しくは有機酸のアンモニウム塩、ペプトン、肉エキス若しくはコーンステープリカー、又はその他の含窒素化合物が挙げられる。無機物としては、リン酸第一カリウム、リン酸第二カリウム、リン酸マグネシウム、硫酸マグネシウム、塩化ナトリウム、硫酸第一鉄、硫酸マンガン、硫酸銅、炭酸カルシウム等が挙げられる。培養は、通常、振盪培養又は通気攪拌培養などの好氣的条件下、37℃で2～16時間行う。pHの調整は、無機又は有機酸、アルカリ溶液等を用いて行う。培養中は必要に応じてアンピシリンやテトラサイクリン等の抗生物質を培地に添加してもよい。

【0059】

プロモーターとして誘導性のプロモーターを用いた発現ベクターで形質転換した微生物を培養する場合は、必要に応じてインデューサーを培地に添加してもよい。例えば、イソプロピル-β-D-チオガラクトシド(IPTG)で誘導可能なtacプロ

モーターを有する発現ベクターが組み込まれた宿主を培養するときには IPTG等を培地に添加することができる。また、インドール酢酸(IAA)で誘導可能なtrpプロモーターを用いた発現ベクターで形質転換した微生物を培養するときにはIAA等を培地に添加することができる。

【 0 0 6 0 】

動物細胞を宿主として得られた形質転換体を培養する培地としては、一般に使用されているRPMI1640培地、DMEM培地又はこれらの培地に牛胎児血清等を添加した培地等が挙げられる。培養は、通常、5 %CO₂存在下、37℃で1～30日行う。培養中は必要に応じてカナマイシン、ペニシリン等の抗生物質を培地に添加してもよい。

【 0 0 6 1 】

培養後、本発明のタンパク質が菌体内又は細胞内に生産される場合には、超音波処理、凍結融解の繰り返し、ホモジナイザー処理などを施して菌体又は細胞を破碎することにより目的のタンパク質採取する。また、本発明のタンパク質が菌体外又は細胞外に生産される場合には、培養液をそのまま使用するか、遠心分離等により菌体又は細胞を除去する。その後、タンパク質の単離精製に用いられる一般的な生化学的方法、例えば硫酸アンモニウム沈殿、ゲルクロマトグラフィー、イオン交換クロマトグラフィー、アフィニティークロマトグラフィー等を単独で又は適宜組み合わせて用いることにより、前記培養物中から本発明のタンパク質を単離精製することができる。

【 0 0 6 2 】

5. 修復遺伝子破壊株

本発明において、前記修復酵素によりDNA配列の損傷やミスマッチが修復された遺伝子を修復遺伝子という。そのような修復遺伝子中にマーカーなどの修飾遺伝子などを組み込んだプラスミドを作製し、これを宿主に導入すると、宿主のゲノムにおける相同組換えにより、本来の修復遺伝子が機能を果たさなくなる。このような破壊株においては突然変異の頻度が高くなるため、宿主が有する種々の遺伝子に簡単に突然変異を起こさせることが可能となる。

【 0 0 6 3 】

宿主は特に限定されるものではなく、好熱菌、大腸菌、枯草菌、酵母など遺伝子操作が確立した微生物が挙げられる。

修復遺伝子へのマーカーの組み込みは、制限酵素による切断及びライゲーションにより行う。マーカーとしては、耐熱化したカナマイシン耐性遺伝子等が挙げられるが、これらに限定されるものではない。

【 0 0 6 4 】

6. DNAのエラー配列の修復又はエラー合成の防止

PCR (polymerase chain reaction)などにおいて忠実にDNA合成を行う確率があり高くないDNAポリメラーゼ（例えばTaq DNAポリメラーゼ等）を用いると、野生型DNAを増幅させたときの増幅DNA断片に塩基の置換変異を生じることがある。この置換変異を「DNAエラー」といい、PCRにより生じたDNAエラーを「PCRエラー」という。本発明のタンパク質は、このようなDNAエラーを修復する機能を有する。

【 0 0 6 5 】

DNA合成反応（PCR、あるいは各種DNAポリメラーゼによる温度変化を伴わない合成反応など）の組成液中に本発明のタンパク質を混合して反応させることで、既にエラーが生じた配列の修復を行い、また、新たにエラーが生じないように、エラーの防止を図ることができる。

【 0 0 6 6 】

【実施例】

以下、実施例により本発明をさらに具体的に説明する。但し、本発明はこれら実施例にその技術的範囲が限定されるものではない。

〔実施例1〕 *Thermus thermophilus* HB8由来MutY蛋白質の調製

Thermus thermophilus HB8 (ATCC 27634) を *Thermus* 栄養培地 (0.4% トリプトン (DIFCO laboratories)、0.2% 酵母エキス (オリエンタル酵母)、0.1% NaCl, pH 7.5) 5ml に植菌し、70℃、15時間培養した後、5,000rpm、10分の遠心により菌体を回収した。得られた菌体を10mM トリス塩酸 (pH 8.0)、10mM EDTA 500 μ l に懸濁し、10mgの卵白リゾチーム (生化学工業株式会社) を加えて42℃で20分間インキュベートし溶菌した。100℃、10分間の熱処理を行ったRnaseA (Sigma社) 溶液

を最終濃度が0.1mg/mlとなるように加えて37℃、30分間反応させた後、100 μ lの10% SDS溶液を加えて37℃、20分間インキュベートした。1mgのプロテアーゼKを加え、50℃で30分間反応させ、さらに65℃で15分間インキュベートした。フェノール抽出、フェノール-クロロフォルム抽出、クロロフォルム抽出を1回ずつ行った後、水層に1.5mlの100%冷エタノールを重層し、析出されたDNAを、先端を丸めたパスツールピペットで巻き取った。70%エタノール、100%エタノールによりDNAをリンスし、乾燥させた後、一晩かけて200 μ lの10mM トリス塩酸 (pH8.0)、10mM EDTA (TE) 緩衝液に溶解した。

【 0 0 6 7 】

下記の5' 側プライマー及び3' 側プライマーを用い、*Thermus thermophilus* HB8 ゲノムDNAを鋳型としてPCRを行った。

5' 側プライマー：5' -ATATCATATggaAgCCTggCggAAAagCCCTCCTCgCCT-3' (配列番号9)

3' 側プライマー：5' -ATATAgATCTTTATTATgCgTCCgggAggggggACTACgCCC-3' (配列番号10)

【 0 0 6 8 】

反応混合液を10分間煮沸後、2.5ユニットのTaqDNAポリメラーゼを添加し、Program Temp Control SystemPC-700 (アステック社) を用いてPCRを行い、MutYをコードするDNAフラグメントを増幅した。増幅反応条件は95℃で3分間加熱後、95℃-1分、55℃-3分、72℃-1分のサイクルを30回繰り返し、最後に72℃-5分で伸長を完全に停止した。このPCR産物とベクターpT7 Blue(Novagen)をTAクローニング法の原理でライゲーション反応により接続し、得られた組換えベクターpT7Blue-MutYを、100mM NaCl、10mM MgCl₂、1mM ジチオトレイトール (DTT)、50mM Tris-HClバッファー (pH7.5) 40 μ l中にて制限酵素NdeI 2ユニット、BglII 1ユニットにより37℃で1時間分解した。制限酵素NdeI、BamHIによる分解、及びバクテリア由来アルカリフォスファターゼによる末端リン酸基の除去を行ったベクターpET-3aと、上述の処理を行ったPCR産物とをライゲーション反応により接続し、得られた組換えベクターpET-3a-MutYを用いて大腸菌BL21 (DE3) pLysEを形質転換した。

【 0 0 6 9 】

なお、PRISM シークエンシングキット (Perkin Elmer社) を用いて反応を行い、ABI377 (Applied Biosystems社) を用いて解析することにより MutY をコードする遺伝子の塩基配列を決定した。その結果、配列番号 1 に示される塩基配列が得られた。また、該遺伝子によりコードされるアミノ酸配列を配列番号 2 に示す。

【 0 0 7 0 】

上記形質転換体を LBamp 培地 (1 % バクトトリプトン、0.5 % イーストエクストラクト、1 % NaCl、50 μ g/ml アンピシリン) 2ml に植菌し、37°C で 16 時間培養した。得られた培養液を LBamp 培地 1 L に加え、37°C で 3 ~ 4 時間培養し、対数増殖期に達したところで 50 μ g/ml イソプロピル- β -D-ガラクトシド (IPTG) を加え、さらに 5 ~ 6 時間培養した。遠心により菌体を回収し、TE により洗浄した後、20mM トリス塩酸、0.2M NaCl、5mM イミダゾール、1mM 2-メルカプトエタノール、pH8.0、20ml に懸濁し、超音波処理により菌体を破碎した。この破碎液を 10,000 g で 30 分間遠心し、上清を得た。

【 0 0 7 1 】

得られた上清中の MutY 蛋白質をフォスフォセルロースカラム、ハイドロキシアパタイトカラム、フェニルトヨパールカラム (東ソー)、MonoS カラムを用いて単離精製した。MonoS カラムにより、MutY タンパク質が単離されたことを確認した。得られた MutY 蛋白質の純度は SDS ポリアクリルアミドゲル電気泳動 (12.5 % ゲル) により確認した (図 3)。図 3 において、各レーンは以下の通りである。

【 0 0 7 2 】

- レーン 1 : 溶菌粗液
- レーン 2 : フォスフォセルロースカラム画分
- レーン 3 : ハイドロキシアパタイトカラム画分
- レーン 4 : フェニルトヨパールカラム画分
- レーン 5 : MonoS カラム画分

【 0 0 7 3 】

〔実施例 2〕 MutY の理化学的性質

(1) 吸収スペクトル

50mM リン酸カリウム pH7.5、0.8M KCl、1mM DTT、1mM EDTA及び10%グリセロールを含む溶液中において測定した。

結果を図8に示す。図8より、410nm付近に鉄硫黄クラスターに特有なスペクトルを有することが分かった。

【0074】

(2) CDスペクトル

CDスペクトル：50mM Tris-HCl pH8.0、0.1M KCl、1mM DTE、1mM EDTA及び20%グリセロールを含む溶液中において測定した。

その結果、 α -ヘリックス含有率は約41%であった（図9）。

【0075】

(3) 熱安定性

50mM リン酸カリウム pH7.5、0.1M KCl、1mM DTE、1mM EDTA及び20%グリセロールを含む溶液中において測定温度を変化させてCDスペクトルを測定した。

その結果、中性条件（pH7.5）で24℃～80℃で（特に75℃まで）安定であった（図10）。

【0076】

〔実施例3〕 MutY蛋白質による二本鎖DNA中のミスマッチ部位の検出

ミスマッチ部位の検出は、以下の手法により行った（図6）

1箇所のミスマッチ部位を有する基質二本鎖DNAを合成した。これらの二本鎖のうち、ミスマッチを有する方の鎖の5'末端は放射性リン酸基で標識されている（図6、11）。MutY蛋白質100nMを含む反応液（20mMトリス塩酸、100mM KCl、10mM EDTA、10%グリセロール、pH7.5、）に基質DNA 20nMを加え、25℃又は50℃にて30分反応させた。反応液の一部に最終濃度0.2NのNaOHを加えた。NaOHを加えたもの、加えないものそれぞれを、7M尿素含有変性ポリアクリルアミドゲル電気泳動（15%ゲル）にかけた。泳動終了後のゲルについてオートラジオグラフィを行い、反応産物を解析した（図7）。図7において、各レーンと処理条件との関係は表1の通りである。その結果、A:G（オキソグアニン）、A:G、G:Gのミスマッチを有する基質DNAについては低分子の反応産物が確認された。

【0077】

【表1】

基質DNA	A:G	A:T	A:GO	A:C	A:G	G:GO
MutY	-	+	+	+	+	+
NaOH処理	+	-	-	-	-	-
温度(°C)	25	25	25	25	25	25
レーン番号	1	2	3	4	5	6
		7	8	9	10	11
		12	13	14	15	16
		17	18	19	20	21

【0078】

この結果は、MutYがこれらのミスマッチを検出し、そのN-グリコシラーゼ活性およびAPリアーゼ活性により、基質DNAをミスマッチ部位で切断したことを示す。

【 0 0 7 9 】

〔実施例 4〕 *Thermus thermophilus* HB8由来RecJ遺伝子産物の調製

下記の 5' 側プライマー及び 3' 側プライマーを用いたこと以外は実施例 1 と同様にして、*Thermus thermophilus* HB8ゲノムDNAを鋳型としてPCRを行った。

【 0 0 8 0 】

5' 側プライマー：5'-ATCATATgAgAgACCgggTCCgCTggCgggT-3' (配列番号11)

3' 側プライマーとして：5'-ATAgATCTTTACAggTCCACCgCCTggACCTC-3' (配列番号12)

【 0 0 8 1 】

制限酵素NdeI、BamHIによる分解およびバクテリア由来アルカリフォスファターゼによる末端リン酸基の除去を行ったベクターpET-19bと、上述の処理を行ったPCR産物とをライゲーション反応により接続し、得られた組換えベクターpET-19b-RecJを用いて大腸菌BL21(DE3)を形質転換した。

なお、実施例 1 と同様にしてRecJをコードする遺伝子の塩基配列を決定した結果、配列番号 3 に示される塩基配列が得られた。また、該遺伝子によりコードされるアミノ酸配列を配列番号 4 に示す。

【 0 0 8 2 】

上記形質転換体をLBamp培地2mlに植菌し、37℃で16時間培養した。得られた培養液をLBamp培地1Lに加え、37℃で3～4時間培養し、対数増殖期に達したところで50 μ g/mlイソプロピル- β -D-ガラクトシド (IPTG) を加え、さらに5～6時間培養した。遠心により菌体を回収し、TEにより洗浄した後、吸着バッファー (20mMトリス塩酸、0.2M NaCl、5mMイミダゾール、1mM 2-メルカプトエタノール、pH8.0) 20mlに懸濁し、超音波処理により菌体を破碎した。この破碎液を10,000 gで30分間遠心し、沈殿を得た。

【 0 0 8 3 】

得られた沈殿を、6M尿素を含む吸着バッファーに溶解した。この溶解液中の

ヒスチジンタグ付RecJ蛋白質をキレーティングセファロースに吸着させた。すなわち、ニッケルイオンを結合させた後、6M尿素を含む吸着バッファーで十分に洗浄したキレーティングセファロースに、沈殿の溶解液を加え、4℃、1時間インキュベートした。遠心によりセファロース担体を回収し、吸着バッファーで十分に洗浄した。この後、尿素濃度を4M、3M、2M、1Mと段階的に下げた吸着バッファーにより洗浄し、ヒスチジンタグ付RecJ蛋白質のリフォールディングを行った。溶出バッファー（20mMトリス塩酸、0.5M NaCl、500mMイミダゾール、1mM 2-メルカプトエタノール、pH8.0）によりヒスチジンタグ付RecJ蛋白質を溶出した。ヒスチジンタグ付RecJの純度は12.5%SDSポリアクリルアミドゲル電気泳動により確認した（図13）。図13において、各レーンは以下の通りである。

【0084】

M:分子量マーカー

レーン1：全細胞溶解物

レーン2：細胞溶解物（上清）

レーン3：細胞溶解物（ペレット）

レーン4：Ni-NTAカラムによるクロマトグラフィー画分

レーン5：リフォールディング

レーン6：MonoQカラムによる陰イオン交換クロマトグラフィー画分

【0085】

矢印部分は、ヒスチジンタグ付RecJ蛋白質 $[(\text{His})_{10}\text{-RecJ}]$ を意味する。

精製したヒスチジンタグ付RecJ蛋白質を、100ユニットのサーモリシン（Sigma）を用い、25℃で6時間の限定分解を行って、分子量45kdの可溶性コアドメインを得た。

【0086】

【実施例5】 *Thermus thermophilus* HB8由来RecJ蛋白質の理化学的性質

(1) CDスペクトル

50mM リン酸カリウム、100mM KCl、0.1mM DTE及び0.1mM EDTA(pH7.2)を含む溶液中のRecJ 1.6 μM についてCDスペクトルを測定した。測定結果を図15に示す。図15より、約50%の α ヘリックスを含有していた。

【 0 0 8 7 】

(2) 熱安定性試験

実施例4で得られたコアドメイン1.6 μ Mを用い、100mM KCl、0.1mM ジチオスレイトール、0.1mM EDTA、50mMリン酸カリウムバッファー (pH7.5) 中にて、測定温度を変化させてCDスペクトルを解析することにより熱安定性を調べた。この結果、RecJ蛋白質のコアドメインは、15℃～60℃で安定であることがわかった (図16)。

【 0 0 8 8 】

【実施例6】 *Thermus thermophilus* HB8由来RecJ蛋白質のエキソヌクレアーゼ活性測定

実施例4で得られたヒスチジンタグ付RecJ蛋白質をトロンピンで分解することによりタグ部分を除いたもの0.1mMを含む反応液 (20mM Tris-HCl、10mM MgCl₂、100mM KCl、1mM DTT、pH7.5) に5' 末端を放射性リン酸基で標識した49merの一本鎖DNA (下記) を基質として加え、25℃、37℃又は50℃で反応させた (図17)。

【 0 0 8 9 】

一本鎖DNA : 5'-ACTACTTggTACACTgACgCgAgCACgCaggAgCTCATTCCAgtgCGCA-3' (配列番号13)

反応産物を、ポリアクリルアミドゲル電気泳動にかけることにより解析した。この結果、時間の経過にしたがって基質の減少および遊離した放射性リン酸標識ヌクレオチド量の増加が確認され、RecJ蛋白質が5' → 3' 方向のエキソヌクレアーゼ活性を有することが示された (図18)。

図18はRecJの5' → 3' エキソヌクレアーゼ活性を示す。図19はエキソヌクレアーゼ活性のRecJ濃度依存性の結果を示す。RecJのエキソヌクレアーゼ活性は、RecJ濃度に依存して上昇した。また、その活性は高温 (50℃) でさらに上昇した。

【 0 0 9 0 】

図20はエテノヌクレオチドのRecJのエキソヌクレアーゼ活性に与える影響を調べた結果を示す図である。エテノヌクレオチドとは、蛍光標識されたヌクレオチドであり、DNAに取り込まれている状態よりも遊離した状態のほうが強い蛍光を発

することを特徴とする。蛍光強度を測定することにより、エテノヌクレオチドが遊離したかどうか、すなわちDNAが分解されたか否かを知ることができる。エテノヌクレオチドによって標識されたDNAと通常のDNAの両者に対してRecJのエキソヌクレアーゼ活性はほとんど変わらなかった。従って、エテノヌクレオチド標識DNAがRecJの基質となることが分かった。

【 0 0 9 1 】

次に、 $32\mu\text{M}$ エテノヌクレオチド (ϵDNA)、 $0.4\mu\text{M}$ RecJ、 20mM Tris-HCl、 10mM MgCl_2 、 100mM KCl、 $\text{pH}7.5$ を含む反応組成液を 37°C でインキュベートし、励起波長 305nm により蛍光の検出を行った。

結果を図21に示す(蛍光スペクトルの図)。RecJのエキソヌクレアーゼ活性によってエテノヌクレオチドがDNAから遊離することによって、蛍光強度が増加した。

【 0 0 9 2 】

また、 $32\mu\text{M}$ エテノヌクレオチド (ϵDNA)、 $0.4\mu\text{M}$ RecJ、 20mM Tris-HCl、 10mM MgCl_2 、 100mM KCl、 $\text{pH}7.5$ を含む反応組成液を 37°C でインキュベートし、励起波長 305nm 、蛍光波長 410nm により蛍光強度及び蛍光偏光度の時間変化を測定した。

結果を図22に示す。上パネルは蛍光強度の時間変化を示し、下パネルは蛍光偏光度の時間変化を示す。 ϵDNA にRecJを反応させると、蛍光物質の自由度を示す蛍光偏光度が上昇した。このことは、エテノヌクレオチドのDNAからの遊離を示すと考えられる。

【 0 0 9 3 】

さらに、 $0.1\mu\text{M}$ RecJ、 20mM Tris-HCl、 10mM MgCl_2 、 100mM KCl、 $\text{pH}7.5$ を含む反応組成液を 37°C でインキュベートし、励起波長 305nm 、蛍光波長 410nm により蛍光を検出してエキソヌクレアーゼ活性のDNA濃度依存性を測定した。

その結果を図23に示す。ミカエリス-メンテンの式に従って速度論的パラメータを求めた結果、 $k_{\text{cat}}=0.034/\text{秒}$ 、 $K_{\text{m}}=6.2\mu\text{M}$ であった。

【 0 0 9 4 】

〔実施例7〕 *Thermus thermophilus* HB8由来RecF遺伝子産物の調製

下記の 5' 側プライマー及び 3' 側プライマーを用いたこと以外は実施例 1 と同様に、*Thermus thermophilus* HB8ゲノムDNAを鋳型としてPCRを行った。

【 0 0 9 5 】

5' 側プライマー : 5'-ATATCATATgCgTCTTCTCCTCTTCCggCAACggAACT-3' (配列番号14)

3' 側プライマー : 5'-ATATAgATCTTTATTAggCgCCAgggCACAggACCACCCCT-3' (配列番号15)

【 0 0 9 6 】

制限酵素NdeI、BamHIによる分解およびバクテリア由来アルカリフォスファターゼによる末端リン酸基の除去を行ったベクターpET-15bと、上述の処理を行ったPCR産物をライゲーション反応により接続し、得られた組換えベクターpET-15b-RecFを用いて大腸菌BL21(DE3)pLysEを形質転換した。

なお、実施例 1 と同様にしてRecFをコードする遺伝子の塩基配列を決定した結果、配列番号 5 に示される塩基配列が得られた。また、該遺伝子によりコードされるアミノ酸配列を配列番号 6 に示す。

【 0 0 9 7 】

上記形質転換体をLBamp培地2mlに植菌し、37℃で16時間培養した。得られた培養液をLBamp培地1Lに加え、37℃で 3 ～ 4 時間培養し、対数増殖期に達したところで50 μ g/mlイソプロピル-1-チオ- β -D-ガラクトシド (IPTG) を加え、さらに 5 ～ 6 時間培養した。遠心により菌体を回収し、TEにより洗浄した後、吸着バッファー (20mMトリス塩酸、0.2M NaCl、5mMイミダゾール、1mM 2-メルカプトエタノール、pH8.0) 20mlに懸濁し、超音波処理により菌体を破碎した。この破碎液を10,000 gで30分間遠心し、上清を得た。

【 0 0 9 8 】

得られた上清中のヒスチジンタグ付RecF蛋白質をキレーティングセファロースカラムに吸着させた。すなわち、ニッケルイオンを結合させた後、吸着バッファーで平衡化したキレーティングセファロースカラムに上清をかけ、さらに吸着バッファーによるカラムの洗浄を行った。この後、溶出バッファー (20mMトリス塩酸、0.2M NaCl、500mMイミダゾール、1mM 2-メルカプトエタノール、pH8.0) に

よりヒスチジンタグ付RecF蛋白質を溶出した。ヒスチジンタグ付RecF蛋白質の純度は12.5%SDSポリアクリルアミドゲル電気泳動により確認した(図25)。

図25において、各レーンは以下の通りである。

【 0 0 9 9 】

M:分子量マーカー

T:全細胞溶解物

S:細胞溶解物(上清)

His:ヒスチジン結合タンパク質

HA:ヒドロキシアパタイトカラムクロマトグラフィー画分

【 0 1 0 0 】

〔実施例8〕 RecFの理化学的性質

(1) CDスペクトル

50mM Tris-HCl、100mM KCl (pH7.5)を含む溶液中、1.4 μ MのRecFについてCDスペクトルを測定した。その結果、RecFは約40%の α ヘリックスを含有していた。

【 0 1 0 1 】

(2) 熱安定性試験

50mM Tris-HCl、100mM KCl(pH7.5)を含む溶液中、測定温度を変化させてCDスペクトルを解析することにより熱安定性を調べた。

その結果、RecFは、50℃まで安定であることがわかった。

【 0 1 0 2 】

(3) 結合作用解析

5 μ M RecF、50mM Tris-HCl、100mM KCl、0.1mM EDTA、5mM 2-メルカプトエタノール、pH7.5を含む反応組成液及びエテノDNAを25℃でインキュベートし、励起波長310nmにより蛍光スペクトルを解析した。 ϵ DNA存在下でRecFのスペクトルに変化が見られたことから、RecFはDNAと結合するものであり、解離定数は5.3 μ Mであった(図28)。

【 0 1 0 3 】

(4) ATPase活性

1 μ M RecF、50mM Tris-HCl、pH7.5、10mM 酢酸マグネシウム、100mM KCl、2mM

ホスホエノールピルビン酸、0.3mM NADH、1mM DTE、25U ピルビン酸キナーゼ、25U 乳酸デヒドロゲナーゼを含む反応組成液をインキュベートし、ATPase活性を測定した。その結果、RecF単独でもATPase活性を有し、温度を25℃から37℃に上げると活性が上昇した(図29)。

【0104】

また、1μM RecF、50mM Tris-HCl、pH7.5、6μM poly(dT)又は6μM poly(dA)・poly(dT)、10mM 酢酸マグネシウム、100mM KCl、2mM ホスホエノールピルビン酸、0.3mM NADH、1mM DTE、25U ピルビン酸キナーゼ、25U 乳酸デヒドロゲナーゼを含む反応組成液を25℃でインキュベートし、ATPase活性を測定した。その結果、一本鎖DNA (poly(dT)) 存在下では活性が増加し、二本鎖DNA (poly(dA)・poly(dT)) 存在下では活性が減少した(図30)。

【0105】

〔実施例9〕 *Thermus thermophilus* HB8由来TRCF (transcription-repair coupling factor) 遺伝子産物の調製

下記の5'側プライマー及び3'側プライマーを用いたこと以外は実施例1と同様にして、*Thermus thermophilus* HB8ゲノムDNAを鋳型としてPCRを行った。

【0106】

5'側プライマー：5'-ATATCATATggAAATCgCgCTAgAgAggATCTACggCC-3' (配列番号16)

3'側プライマー：5'-ATATAgATCTTTATTAGAGGTCGGCGAAGAGGTAGAGCACC-3' (配列番号17)

【0107】

制限酵素NdeI、BamHIによる分解およびバクテリア由来アルカリフォスファターゼによる末端リン酸基の除去を行ったベクターpET-15bと、上述の処理を行ったPCR産物をライゲーション反応により接続し、得られた組換えベクターpET-15b-TRCFを用いて大腸菌BL21(DE3)pLysEを形質転換した。

なお、実施例1と同様にしてTRCFをコードする遺伝子の塩基配列を決定した結果、配列番号7に示される塩基配列が得られた。また、該遺伝子によりコードされるアミノ酸配列を配列番号8に示す。

【 0 1 0 8 】

上記形質転換体をLBamp培地2mlに植菌し、37℃で16時間培養した。得られた培養液をLBamp培地1Lに加え、37℃で3～4時間培養し、対数増殖期に達したところで50 μ g/mlイソプロピル-1-チオ- β -D-ガラクトシド (IPTG) を加え、さらに5～6時間培養した。遠心により菌体を回収し、TEにより洗浄した後、吸着バッファー (20mMトリス塩酸、0.2M NaCl、5mMイミダゾール、1mM 2-メルカプトエタノール、pH8.0) 20mlに懸濁し、超音波処理により菌体を破碎した。この破碎液を10,000 gで30分間遠心し、上清を得た。

【 0 1 0 9 】

得られた上清中のヒスチジンタグ付TRCF蛋白質をキレーティングセファロースカラムに吸着させた。すなわち、ニッケルイオンを結合させた後、吸着バッファーで平衡化したキレーティングセファロースカラムに上清をかけ、さらに吸着バッファーによるカラムの洗浄を行った。この後、溶出バッファー (20mMトリス塩酸、0.2M NaCl、500mMイミダゾール、1mM 2-メルカプトエタノール、pH8.0) によりヒスチジンタグ付RecF蛋白質を溶出した。ヒスチジンタグ付TRCF蛋白質の純度は12.5%SDSポリアクリルアミドゲル電気泳動により確認した (図33)。図33において、上パネルはUvrB- β の精製結果を示し、下パネルはTRCF- β の精製結果を示す。上パネルの各レーンは以下の通りである。

【 0 1 1 0 】

M:分子量マーカー

- ①: 全細胞溶解物
- ②: 細胞溶解物遠心上清
- ③: ニッケルカラムクロマトグラフィー画分
- ④: ブチルカラムクロマトグラフィー画分

下パネルの各レーンは以下の通りである。

M: 分子量マーカー

- ①: 全細胞溶解物
- ②: ニッケルカラム及びブチルカラムクロマトグラフィー画分

【 0 1 1 1 】

〔実施例10〕 TRCFの理化学的性質

(1) CDスペクトル

50mM Tris-HCl、100mM KCl (pH7.9)を含む溶液中、25℃におけるUvrB- β 及びTRCFについてCDスペクトルを測定した。結果を図35に示す。UvrB- β 及びTRCFは、類似した立体構造を有することが分かった。

【0112】

(2) 熱安定性試験

50mM Tris-HCl、100mM KCl (pH7.9)を含む溶液中、測定温度を変化させてUvrB- β 及びTRCF- β のCDスペクトルを解析することにより熱安定性を調べた。

その結果、UvrB- β 及びTRCF- β は、ともにpH7.9において20℃～75℃で安定であることがわかった (図36)。

【0113】

(3) pH安定性試験

100mM KClを含み、pHの異なる各種緩衝液中において、UvrB- β 及びTRCF- β のCDスペクトルを測定した。

その結果、TRCFは、25℃においてpH4から9まで安定であることがわかった (図37)。

【0114】

(4) 結合作用解析

センサーチップNTAにNiCl₂をインジェクトした後、 β ドメインをインジェクトし、固定化した。UvrA及びUvrBは、ATP存在下でのみ相互作用することが知られているので、各 β ドメインとUvrAとの相互作用も、ATP存在下及び非存在下の両方で測定した (図38)。

その結果、解離定数(K_d)は、ATP存在下では0.5 μ M、ATP非存在下では1.3 μ Mであった (図39)。

【0115】

【発明の効果】

本発明により、DNA修復酵素及びそれをコードする遺伝子が提供される。本発明の酵素は、DNA修復活性を有し、かつ温度安定性に優れているため、生化学、

分子生物学等の研究用試薬、及び各種DNA合成反応におけるエラーの防止用又は修復用試薬として有用である。

【 0 1 1 6 】

【配列表】

SEQUENCE LISTING

<110> RIKEN

<120> GENE ENCODING DNA REPAIR ENZYME

<130> RJH12-082T

<140>

<141>

<160> 17

<170> PatentIn Ver. 2.0

<210> 1

<211> 975

<212> DNA

<213> *Thermus thermophilus*

<220>

<221> CDS

<222> (1)..(975)

<400> 1

gtg gag gcc tgg cgg aaa gcc ctc ctc gcc tgg tac cgg gaa aac gcc 48

Val Glu Ala Trp Arg Lys Ala Leu Leu Ala Trp Tyr Arg Glu Asn Ala

1 5 10 15

cgc ccc ctc ccc tgg cgg ggg gag aag gac cct tac cgc gtc ctg gtc 96

Arg Pro Leu Pro Trp Arg Gly Glu Lys Asp Pro Tyr Arg Val Leu Val

20 25 30

tcc gag gtc ctt ctg cag cag acc cgg gtg gag cag gcc ctc ccc tat 144

Ser Glu Val Leu Leu Gln Gln Thr Arg Val Glu Gln Ala Leu Pro Tyr

35 40 45

tac cgc cgc ttt ctg gag cgc ttt ccc acc ctg aag gcc ctg gcc gcg 192

Tyr Arg Arg Phe Leu Glu Arg Phe Pro Thr Leu Lys Ala Leu Ala Ala

50 55 60

gct tcc ctg gaa gag gtc ctt agg gtc tgg cag ggg gcg ggc tac tac 240

Ala Ser Leu Glu Glu Val Leu Arg Val Trp Gln Gly Ala Gly Tyr Tyr

65 70 75 80

cgg cgg gcg gaa cac ctc cac cgc ctg gcc cga agc gtg gag gag ctt 288

Arg Arg Ala Glu His Leu His Arg Leu Ala Arg Ser Val Glu Glu Leu

85 90 95

ccc ccg agc ttc gcc gag ctt cgg ggg ctt cct ggt ctc ggg cct tac 336

Pro Pro Ser Phe Ala Glu Leu Arg Gly Leu Pro Gly Leu Gly Pro Tyr

100 105 110

acc gcg gcg gcg gtg gcc tcc atc gcc ttc ggg gag cgg gtg gcg gcg 384

Thr Ala Ala Ala Val Ala Ser Ile Ala Phe Gly Glu Arg Val Ala Ala

115	120	125	
gtg gac ggg aac gtc cgg agg gtc ctc tcc cgc ctc ttc gcc cgg gaa			432
Val Asp Gly Asn Val Arg Arg Val Leu Ser Arg Leu Phe Ala Arg Glu			
130	135	140	
agc ccc aag gag aag gag ctt ttc gcc ctc gcc cag ggc ctc ctc ccc			480
Ser Pro Lys Glu Lys Glu Leu Phe Ala Leu Ala Gln Gly Leu Leu Pro			
145	150	155	160
gag ggc gtg gac ccg ggg gtg tgg aac cag gcc ctc atg gag ctc ggg			528
Glu Gly Val Asp Pro Gly Val Trp Asn Gln Ala Leu Met Glu Leu Gly			
165	170	175	
gcc acg gtc tgc ctg ccg aaa cgg ccc cgt tgc ggg gcc tgc ccc cta			576
Ala Thr Val Cys Leu Pro Lys Arg Pro Arg Cys Gly Ala Cys Pro Leu			
180	185	190	
ggg gcc ttc tgc cgg ggg aag gag gcc ccc ggg cgc tac ccc gcg ccc			624
Gly Ala Phe Cys Arg Gly Lys Glu Ala Pro Gly Arg Tyr Pro Ala Pro			
195	200	205	
agg aag cgc cgg gcg aag gag gag cgc ctc gtc gcc ctc gtc ctc ctc			672
Arg Lys Arg Arg Ala Lys Glu Glu Arg Leu Val Ala Leu Val Leu Leu			
210	215	220	
ggg cgg aag ggg gtg cac ctg gaa agg ctt gag ggg cgc ttc cag ggc			720
Gly Arg Lys Gly Val His Leu Glu Arg Leu Glu Gly Arg Phe Gln Gly			
225	230	235	240

ctc tac ggc gtc ccc ctc ttt ccc cct gag gag ctt ccc ggg cgg gag 768

Leu Tyr Gly Val Pro Leu Phe Pro Pro Glu Glu Leu Pro Gly Arg Glu

245

250

255

gcg gcc ttc ggg gtg agg tct agg ccc cta ggc gag gtg cgc cac gcc 816

Ala Ala Phe Gly Val Arg Ser Arg Pro Leu Gly Glu Val Arg His Ala

260

265

270

ctc acc cac cgg agg ctt cgc gtg gag gtg cgg ggg gcc ctt tgg gaa 864

Leu Thr His Arg Arg Leu Arg Val Glu Val Arg Gly Ala Leu Trp Glu

275

280

285

ggg gag ggg gag gac ccc tgg aag agg ccc cta ccc aag ctc atg gag 912

Gly Glu Gly Glu Asp Pro Trp Lys Arg Pro Leu Pro Lys Leu Met Glu

290

295

300

aag gtg ctc cgc aag gcg ctt ccc ctc ctc gct cat gcg ggc gta gtc 960

Lys Val Leu Arg Lys Ala Leu Pro Leu Leu Ala His Ala Gly Val Val

305

310

315

320

ccc ctc ccg gac gca

975

Pro Leu Pro Asp Ala

325

<210> 2

<211> 325

<212> PRT

<213> *Thermus thermophilus*

<400> 2

Val Glu Ala Trp Arg Lys Ala Leu Leu Ala Trp Tyr Arg Glu Asn Ala

1

5

10

15

Arg Pro Leu Pro Trp Arg Gly Glu Lys Asp Pro Tyr Arg Val Leu Val

20

25

30

Ser Glu Val Leu Leu Gln Gln Thr Arg Val Glu Gln Ala Leu Pro Tyr

35

40

45

Tyr Arg Arg Phe Leu Glu Arg Phe Pro Thr Leu Lys Ala Leu Ala Ala

50

55

60

Ala Ser Leu Glu Glu Val Leu Arg Val Trp Gln Gly Ala Gly Tyr Tyr

65

70

75

80

Arg Arg Ala Glu His Leu His Arg Leu Ala Arg Ser Val Glu Glu Leu

85

90

95

Pro Pro Ser Phe Ala Glu Leu Arg Gly Leu Pro Gly Leu Gly Pro Tyr

100

105

110

Thr Ala Ala Ala Val Ala Ser Ile Ala Phe Gly Glu Arg Val Ala Ala

115

120

125

Val Asp Gly Asn Val Arg Arg Val Leu Ser Arg Leu Phe Ala Arg Glu

130

135

140

Ser Pro Lys Glu Lys Glu Leu Phe Ala Leu Ala Gln Gly Leu Leu Pro

145 150 155 160

Glu Gly Val Asp Pro Gly Val Trp Asn Gln Ala Leu Met Glu Leu Gly

165 170 175

Ala Thr Val Cys Leu Pro Lys Arg Pro Arg Cys Gly Ala Cys Pro Leu

180 185 190

Gly Ala Phe Cys Arg Gly Lys Glu Ala Pro Gly Arg Tyr Pro Ala Pro

195 200 205

Arg Lys Arg Arg Ala Lys Glu Glu Arg Leu Val Ala Leu Val Leu Leu

210 215 220

Gly Arg Lys Gly Val His Leu Glu Arg Leu Glu Gly Arg Phe Gln Gly

225 230 235 240

Leu Tyr Gly Val Pro Leu Phe Pro Pro Glu Glu Leu Pro Gly Arg Glu

245 250 255

Ala Ala Phe Gly Val Arg Ser Arg Pro Leu Gly Glu Val Arg His Ala

260 265 270

Leu Thr His Arg Arg Leu Arg Val Glu Val Arg Gly Ala Leu Trp Glu

275 280 285

Gly Glu Gly Glu Asp Pro Trp Lys Arg Pro Leu Pro Lys Leu Met Glu

290 295 300

Lys Val Leu Arg Lys Ala Leu Pro Leu Leu Ala His Ala Gly Val Val

305 310 315 320

Pro Leu Pro Asp Ala

325

<210> 3

<211> 1998

<212> DNA

<213> *Thermus thermophilus*

<220>

<221> CDS

<222> (1)..(1998)

<400> 3

atg agg gac cgg gtc cgc tgg cgg gtg ctt tcc ctc cct ccc ctc gcc 48

Met Arg Asp Arg Val Arg Trp Arg Val Leu Ser Leu Pro Pro Leu Ala

1 5 10 15

cag tgg cgg gag gtg atg gcg gcc ttg gag gtg ggg ccg gag gcc gcc 96

Gln Trp Arg Glu Val Met Ala Ala Leu Glu Val Gly Pro Glu Ala Ala

20 25 30

ctg gcc tac tgg cac cgg ggc ttt agg cgc aag gag gac ctg gac ccc 144

Leu Ala Tyr Trp His Arg Gly Phe Arg Arg Lys Glu Asp Leu Asp Pro

35 40 45

ccc ctc gcc ctc ctt ccc ctc aag ggc ctg agg gag gcg gcg gcc ctc 192

Pro Leu Ala Leu Leu Pro Leu Lys Gly Leu Arg Glu Ala Ala Ala Leu

50

55

60

ctg gag gag gcg ctc cgc cag ggg aag cgg atc cgc gtc cac ggg gac 240

Leu Glu Glu Ala Leu Arg Gln Gly Lys Arg Ile Arg Val His Gly Asp

65

70

75

80

tac gac gcc gac ggg ctc acg ggc acg gcc atc ctg gtt cgg ggc ctc 288

Tyr Asp Ala Asp Gly Leu Thr Gly Thr Ala Ile Leu Val Arg Gly Leu

85

90

95

gcc gcc ttg ggc gcc gac gtc cac ccc ttc atc ccc cac cgg ctg gag 336

Ala Ala Leu Gly Ala Asp Val His Pro Phe Ile Pro His Arg Leu Glu

100

105

110

gaa ggg tac ggg gtg ctg atg gag cgg gtt ccc gag cac ctc gag gcc 384

Glu Gly Tyr Gly Val Leu Met Glu Arg Val Pro Glu His Leu Glu Ala

115

120

125

tgc gac ctc ttc ctc acc gtg gac tgc ggg atc acg aac cac gcc gag 432

Ser Asp Leu Phe Leu Thr Val Asp Cys Gly Ile Thr Asn His Ala Glu

130

135

140

ctc agg gag ctt ttg gaa aac ggg gtg gag gtg atc gtc acc gac cac 480

Leu Arg Glu Leu Leu Glu Asn Gly Val Glu Val Ile Val Thr Asp His

145

150

155

160

cac acc ccc ggc aag acc cct tcc ccc ggc ctc gtg gtc cac ccc gcc 528

His Thr Pro Gly Lys Thr Pro Ser Pro Gly Leu Val Val His Pro Ala

165

170

175

ctc acc ccg gac ctt aag gag aag ccc acg ggg gcg ggg gtg gtc ttc 576

Leu Thr Pro Asp Leu Lys Glu Lys Pro Thr Gly Ala Gly Val Val Phe

180

185

190

ctc ctc ctc tgg gcc ctc cac gag cgc ctg ggc ctt ccc cca ccc ctg 624

Leu Leu Leu Trp Ala Leu His Glu Arg Leu Gly Leu Pro Pro Pro Leu

195

200

205

gag tac gcc gac ctc gcc gcg gtg ggc acc atc gcc gac gtg gcc ccc 672

Glu Tyr Ala Asp Leu Ala Ala Val Gly Thr Ile Ala Asp Val Ala Pro

210

215

220

ctt tgg ggc tgg aac cgg gcc ttg gtg aag gag ggc ctg gcc cgc atc 720

Leu Trp Gly Trp Asn Arg Ala Leu Val Lys Glu Gly Leu Ala Arg Ile

225

230

235

240

ccc gcc tcc tcc tgg gtt ggg ctc agg ctt ctg gcc gag gcg gtg ggg 768

Pro Ala Ser Ser Trp Val Gly Leu Arg Leu Leu Ala Glu Ala Val Gly

245

250

255

tac acg ggg aag gcg gtg gag gtg gcc ttc cgc atc gcc ccc cgg atc 816

Tyr Thr Gly Lys Ala Val Glu Val Ala Phe Arg Ile Ala Pro Arg Ile

260

265

270

aac gcg gca agc cgc ctc ggg gag gct gag aag gcc cta agg ctc ctc 864

Asn Ala Ala Ser Arg Leu Gly Glu Ala Glu Lys Ala Leu Arg Leu Leu

275	280	285	
ctc acc gac gac gcg gcc gag gcc cag gcc ctc gtg ggg gaa ctc cac	912		
Leu Thr Asp Asp Ala Ala Glu Ala Gln Ala Leu Val Gly Glu Leu His			
290	295	300	
cgg ctg aac gcc cgc cgc cag acc ctg gag gag gcc atg ctc agg aag	960		
Arg Leu Asn Ala Arg Arg Gln Thr Leu Glu Glu Ala Met Leu Arg Lys			
305	310	315	320
ctc ctc ccc cag gcg gac ccc gag gcc aag gcc atc gtc ctc ctg gac	1008		
Leu Leu Pro Gln Ala Asp Pro Glu Ala Lys Ala Ile Val Leu Leu Asp			
325	330	335	
ccc gag ggg cac ccg ggg gtg atg ggc atc gtg gcg agc cgc atc ctg	1056		
Pro Glu Gly His Pro Gly Val Met Gly Ile Val Ala Ser Arg Ile Leu			
340	345	350	
gag gcc acc ctc cgg ccc gtc ttc ctg gtg gcc cag ggc aag ggg acg	1104		
Glu Ala Thr Leu Arg Pro Val Phe Leu Val Ala Gln Gly Lys Gly Thr			
355	360	365	
gtg cgg agc ctc gcc ccc atc agc gcc gtg gag gcc cta agg agc gcc	1152		
Val Arg Ser Leu Ala Pro Ile Ser Ala Val Glu Ala Leu Arg Ser Ala			
370	375	380	
gag gac ctt ttg ttg cgc tac ggg ggg cac aag gag gcg gcg ggc ttc	1200		
Glu Asp Leu Leu Leu Arg Tyr Gly Gly His Lys Glu Ala Ala Gly Phe			
385	390	395	400

gcc atg gac gag gcc ctc ttc ccc gcc ttc aag gcc cgg gtg gag gcc 1248

Ala Met Asp Glu Ala Leu Phe Pro Ala Phe Lys Ala Arg Val Glu Ala

405

410

415

tac gcc gcc cgc ttc ccc gac ccc gtg cgc gag gtg gcc ctt ttg gac 1296

Tyr Ala Ala Arg Phe Pro Asp Pro Val Arg Glu Val Ala Leu Leu Asp

420

425

430

ctg ctt ccg gag ccc ggc ctc ctc ccc cag gtc ttc cgg gag ctc gcc 1344

Leu Leu Pro Glu Pro Gly Leu Leu Pro Gln Val Phe Arg Glu Leu Ala

435

440

445

ctt ttg gag ccc tac ggc gag gga aac ccc gag ccc ctc ttc ctc ctc 1392

Leu Leu Glu Pro Tyr Gly Glu Gly Asn Pro Glu Pro Leu Phe Leu Leu

450

455

460

ttc ggc gcc ccg gag gag gcc cgg cgc ctc ggg gag ggc cgc cac ctc 1440

Phe Gly Ala Pro Glu Glu Ala Arg Arg Leu Gly Glu Gly Arg His Leu

465

470

475

480

gcc ttc cgc ctg aag ggg gtg cgg gtc ctg gcc tgg aaa cag ggg gac 1488

Ala Phe Arg Leu Lys Gly Val Arg Val Leu Ala Trp Lys Gln Gly Asp

485

490

495

ctc gcc ctg ccc ccg gag gtg gag gtg gcg ggc ctc ctc agc gaa aac 1536

Leu Ala Leu Pro Pro Glu Val Glu Val Ala Gly Leu Leu Ser Glu Asn

500

505

510

gcc tgg aac ggc cac ctc gcc tac gag gtc cag gcg gtg gac ctg cga 1584

Ala Trp Asn Gly His Leu Ala Tyr Glu Val Gln Ala Val Asp Leu Arg

515

520

525

aag cca gag gcg ctg gag ggc ggg atc gcg ccc ttc gcc tac ccc ctg 1632

Lys Pro Glu Ala Leu Glu Gly Gly Ile Ala Pro Phe Ala Tyr Pro Leu

530

535

540

ccc ctc ctc gag gcc ctg gcc cgg gcc cgc ctg ggg gaa ggg gtc tac 1680

Pro Leu Leu Glu Ala Leu Ala Arg Ala Arg Leu Gly Glu Gly Val Tyr

545

550

555

560

gtc ccc gag gac aac cct gag ggg ctg gac tac gcc agg aag gcg ggc 1728

Val Pro Glu Asp Asn Pro Glu Gly Leu Asp Tyr Ala Arg Lys Ala Gly

565

570

575

ttc cgc ctc ctc ccc ccc gag gag gcc ggg ctt tgg ctc ggc ctc ccc 1776

Phe Arg Leu Leu Pro Pro Glu Glu Ala Gly Leu Trp Leu Gly Leu Pro

580

585

590

cca agg ccg gtc ctg ggc agg cgg gtg gag gtg gcc ctg ggg cgg gag 1824

Pro Arg Pro Val Leu Gly Arg Arg Val Glu Val Ala Leu Gly Arg Glu

595

600

605

gcg cgg gcc agg ctt tcc gcc ccc ccc gtc ctc cac acc ccc gag gcc 1872

Ala Arg Ala Arg Leu Ser Ala Pro Pro Val Leu His Thr Pro Glu Ala

610

615

620

cgg ctc aaa gcc ctc gtc cac cgc cgc ctc ctc ttc gcc tac gag cgc 1920

Arg Leu Lys Ala Leu Val His Arg Arg Leu Leu Phe Ala Tyr Glu Arg

625 630 635 640

cgt cac ccg ggc ctc ttc agc gag gcc ctc ctc gcc tac tgg gag gtg 1968

Arg His Pro Gly Leu Phe Ser Glu Ala Leu Leu Ala Tyr Trp Glu Val

645 650 655

aac cgt gta cag gag ccc gcg gga agc cca

1998

Asn Arg Val Gln Glu Pro Ala Gly Ser Pro

660 665

<210> 4

<211> 666

<212> PRT

<213> *Thermus thermophilus*

<400> 4

Met Arg Asp Arg Val Arg Trp Arg Val Leu Ser Leu Pro Pro Leu Ala

1 5 10 15

Gln Trp Arg Glu Val Met Ala Ala Leu Glu Val Gly Pro Glu Ala Ala

20 25 30

Leu Ala Tyr Trp His Arg Gly Phe Arg Arg Lys Glu Asp Leu Asp Pro

35 40 45

Pro Leu Ala Leu Leu Pro Leu Lys Gly Leu Arg Glu Ala Ala Ala Leu

50 55 60

Leu Glu Glu Ala Leu Arg Gln Gly Lys Arg Ile Arg Val His Gly Asp
65 70 75 80

Tyr Asp Ala Asp Gly Leu Thr Gly Thr Ala Ile Leu Val Arg Gly Leu
85 90 95

Ala Ala Leu Gly Ala Asp Val His Pro Phe Ile Pro His Arg Leu Glu
100 105 110

Glu Gly Tyr Gly Val Leu Met Glu Arg Val Pro Glu His Leu Glu Ala
115 120 125

Ser Asp Leu Phe Leu Thr Val Asp Cys Gly Ile Thr Asn His Ala Glu
130 135 140

Leu Arg Glu Leu Leu Glu Asn Gly Val Glu Val Ile Val Thr Asp His
145 150 155 160

His Thr Pro Gly Lys Thr Pro Ser Pro Gly Leu Val Val His Pro Ala
165 170 175

Leu Thr Pro Asp Leu Lys Glu Lys Pro Thr Gly Ala Gly Val Val Phe
180 185 190

Leu Leu Leu Trp Ala Leu His Glu Arg Leu Gly Leu Pro Pro Pro Leu
195 200 205

Glu Tyr Ala Asp Leu Ala Ala Val Gly Thr Ile Ala Asp Val Ala Pro
210 215 220

Leu Trp Gly Trp Asn Arg Ala Leu Val Lys Glu Gly Leu Ala Arg Ile

225 230 235 240

Pro Ala Ser Ser Trp Val Gly Leu Arg Leu Leu Ala Glu Ala Val Gly

245 250 255

Tyr Thr Gly Lys Ala Val Glu Val Ala Phe Arg Ile Ala Pro Arg Ile

260 265 270

Asn Ala Ala Ser Arg Leu Gly Glu Ala Glu Lys Ala Leu Arg Leu Leu

275 280 285

Leu Thr Asp Asp Ala Ala Glu Ala Gln Ala Leu Val Gly Glu Leu His

290 295 300

Arg Leu Asn Ala Arg Arg Gln Thr Leu Glu Glu Ala Met Leu Arg Lys

305 310 315 320

Leu Leu Pro Gln Ala Asp Pro Glu Ala Lys Ala Ile Val Leu Leu Asp

325 330 335

Pro Glu Gly His Pro Gly Val Met Gly Ile Val Ala Ser Arg Ile Leu

340 345 350

Glu Ala Thr Leu Arg Pro Val Phe Leu Val Ala Gln Gly Lys Gly Thr

355 360 365

Val Arg Ser Leu Ala Pro Ile Ser Ala Val Glu Ala Leu Arg Ser Ala

370	375	380	
Glu Asp Leu Leu Leu Arg Tyr Gly Gly His Lys Glu Ala Ala Gly Phe			
385	390	395	400
Ala Met Asp Glu Ala Leu Phe Pro Ala Phe Lys Ala Arg Val Glu Ala			
	405	410	415
Tyr Ala Ala Arg Phe Pro Asp Pro Val Arg Glu Val Ala Leu Leu Asp			
	420	425	430
Leu Leu Pro Glu Pro Gly Leu Leu Pro Gln Val Phe Arg Glu Leu Ala			
	435	440	445
Leu Leu Glu Pro Tyr Gly Glu Gly Asn Pro Glu Pro Leu Phe Leu Leu			
	450	455	460
Phe Gly Ala Pro Glu Glu Ala Arg Arg Leu Gly Glu Gly Arg His Leu			
465	470	475	480
Ala Phe Arg Leu Lys Gly Val Arg Val Leu Ala Trp Lys Gln Gly Asp			
	485	490	495
Leu Ala Leu Pro Pro Glu Val Glu Val Ala Gly Leu Leu Ser Glu Asn			
	500	505	510
Ala Trp Asn Gly His Leu Ala Tyr Glu Val Gln Ala Val Asp Leu Arg			
	515	520	525

Lys Pro Glu Ala Leu Glu Gly Gly Ile Ala Pro Phe Ala Tyr Pro Leu

530

535

540

Pro Leu Leu Glu Ala Leu Ala Arg Ala Arg Leu Gly Glu Gly Val Tyr

545

550

555

560

Val Pro Glu Asp Asn Pro Glu Gly Leu Asp Tyr Ala Arg Lys Ala Gly

565

570

575

Phe Arg Leu Leu Pro Pro Glu Glu Ala Gly Leu Trp Leu Gly Leu Pro

580

585

590

Pro Arg Pro Val Leu Gly Arg Arg Val Glu Val Ala Leu Gly Arg Glu

595

600

605

Ala Arg Ala Arg Leu Ser Ala Pro Pro Val Leu His Thr Pro Glu Ala

610

615

620

Arg Leu Lys Ala Leu Val His Arg Arg Leu Leu Phe Ala Tyr Glu Arg

625

630

635

640

Arg His Pro Gly Leu Phe Ser Glu Ala Leu Leu Ala Tyr Trp Glu Val

645

650

655

Asn Arg Val Gln Glu Pro Ala Gly Ser Pro

660

665

<210> 5

<211> 1029

<212> DNA

<213> *Thermus thermophilus*

<220>

<221> CDS

<222> (1)..(1029)

<400> 5

atg cgg ctt ctc ctc ttc cgg caa cgg aac ttc cgc aac ctg gcc ctg 48

Met Arg Leu Leu Leu Phe Arg Gln Arg Asn Phe Arg Asn Leu Ala Leu

1

5

10

15

gag gcc tac cgc ccc ccg ccg ggc ctt tcc gcc ctg gtg ggg gcc aac 96

Glu Ala Tyr Arg Pro Pro Pro Gly Leu Ser Ala Leu Val Gly Ala Asn

20

25

30

gcc cag ggg aag acg agc ctc ctc ctg ggg atc cac ctg gcc cta ggg 144

Ala Gln Gly Lys Thr Ser Leu Leu Leu Gly Ile His Leu Ala Leu Gly

35

40

45

ggg gag gtc ccc ctg ggc ctt gcc gac ctc gtc cgc ttc ggg gag gag 192

Gly Glu Val Pro Leu Gly Leu Ala Asp Leu Val Arg Phe Gly Glu Glu

50

55

60

gag gcc tgg ctc cac gcc gag gtg gag acg gag ctc ggg gcc tac cgc 240

Glu Ala Trp Leu His Ala Glu Val Glu Thr Glu Leu Gly Ala Tyr Arg

65

70

75

80

ctg gag cac cgc ctg ggc ccc ggg ggg cgg gag gtc ctc ctc aac ggg 288

Leu Glu His Arg Leu Gly Pro Gly Gly Arg Glu Val Leu Leu Asn Gly

85

90

95

aag cgg gtg agc ctt cgg acc ctt tgg gag ctt ccc ggc tgc gtc ctc 336

Lys Arg Val Ser Leu Arg Thr Leu Trp Glu Leu Pro Gly Ser Val Leu

100

105

110

gtc tcc cct ctg gac ctc gag gcg gtc ctc ggg ccc aag gag gag cgg 384

Val Ser Pro Leu Asp Leu Glu Ala Val Leu Gly Pro Lys Glu Glu Arg

115

120

125

cgg gcc tac ctg gac cgg ctc atc gcc cgc ttc tcc cgc cgc tac gcc 432

Arg Ala Tyr Leu Asp Arg Leu Ile Ala Arg Phe Ser Arg Arg Tyr Ala

130

135

140

gcc ctc ctt tcc gcc tac gag aag gcg ctg cgc cag cgg aac gcc ctc 480

Ala Leu Leu Ser Ala Tyr Glu Lys Ala Leu Arg Gln Arg Asn Ala Leu

145

150

155

160

ctc aag gcc ggg ggg gag ggc ctt tcc gcc tgg gac cgg gag ctc gcc 528

Leu Lys Ala Gly Gly Glu Gly Leu Ser Ala Trp Asp Arg Glu Leu Ala

165

170

175

cgc tac ggg gac gag atc gtg gcc ctg agg cgc cgc ttc ctc cgg cgc 576

Arg Tyr Gly Asp Glu Ile Val Ala Leu Arg Arg Arg Phe Leu Arg Arg

180

185

190

ttc gcc ccc atc ctg cgg gag gtc cac gcc gcc ctc gcc gcc aag gag 624

Phe Ala Pro Ile Leu Arg Glu Val His Ala Ala Leu Ala Ala Lys Glu

195	200	205	
gcg ggg ctt cgc ttg gag gag acc gcg ggg gaa ggg gtg ctc cgg gcc	672		
Ala Gly Leu Arg Leu Glu Glu Thr Ala Gly Glu Gly Val Leu Arg Ala			
210	215	220	
ctc gag gcc agc cgg gcc gag gag cgg gaa cgg ggc cag acc ctg gtg	720		
Leu Glu Ala Ser Arg Ala Glu Glu Arg Glu Arg Gly Gln Thr Leu Val			
225	230	235	240
ggg ccc cac cgg gac gac ctg gtc ttc ctc ctg gag ggg cgg ccc gcc	768		
Gly Pro His Arg Asp Asp Leu Val Phe Leu Leu Glu Gly Arg Pro Ala			
245	250	255	
cac cgg ttc gcc agc cgc ggg gag gcc aag acc ctg gcc ctg gcc ctg	816		
His Arg Phe Ala Ser Arg Gly Glu Ala Lys Thr Leu Ala Leu Ala Leu			
260	265	270	
cgc ctc gcc gag cac cgc ctc ctc ggc gag cac cac ggc gag ccc ccc	864		
Arg Leu Ala Glu His Arg Leu Leu Gly Glu His His Gly Glu Pro Pro			
275	280	285	
ctc ctc ctc gtg gac gag tgg ggg gag gag ctg gac gag gcc cgc agg	912		
Leu Leu Leu Val Asp Glu Trp Gly Glu Glu Leu Asp Glu Ala Arg Arg			
290	295	300	
cgg gcc gtc ctc gcc tac gcc cag gcc ctg ccc cag gcc atc ctg gcg	960		
Arg Ala Val Leu Ala Tyr Ala Gln Ala Leu Pro Gln Ala Ile Leu Ala			
305	310	315	320

ggg ctg gaa gcc ccc ccg ggg gtg ccg gta tgc tcg gtg gta cga ggg 1008

Gly Leu Glu Ala Pro Pro Gly Val Pro Val Cys Ser Val Val Arg Gly

325

330

335

gtg gtc ctg tgc cct ggc gcc

1029

Val Val Leu Cys Pro Gly Ala

340

<210> 6

<211> 343

<212> PRT

<213> *Thermus thermophilus*

<400> 6

Met Arg Leu Leu Leu Phe Arg Gln Arg Asn Phe Arg Asn Leu Ala Leu

1

5

10

15

Glu Ala Tyr Arg Pro Pro Pro Gly Leu Ser Ala Leu Val Gly Ala Asn

20

25

30

Ala Gln Gly Lys Thr Ser Leu Leu Leu Gly Ile His Leu Ala Leu Gly

35

40

45

Gly Glu Val Pro Leu Gly Leu Ala Asp Leu Val Arg Phe Gly Glu Glu

50

55

60

Glu Ala Trp Leu His Ala Glu Val Glu Thr Glu Leu Gly Ala Tyr Arg

65

70

75

80

Leu Glu His Arg Leu Gly Pro Gly Gly Arg Glu Val Leu Leu Asn Gly

85

90

95

Lys Arg Val Ser Leu Arg Thr Leu Trp Glu Leu Pro Gly Ser Val Leu

100

105

110

Val Ser Pro Leu Asp Leu Glu Ala Val Leu Gly Pro Lys Glu Glu Arg

115

120

125

Arg Ala Tyr Leu Asp Arg Leu Ile Ala Arg Phe Ser Arg Arg Tyr Ala

130

135

140

Ala Leu Leu Ser Ala Tyr Glu Lys Ala Leu Arg Gln Arg Asn Ala Leu

145

150

155

160

Leu Lys Ala Gly Gly Glu Gly Leu Ser Ala Trp Asp Arg Glu Leu Ala

165

170

175

Arg Tyr Gly Asp Glu Ile Val Ala Leu Arg Arg Arg Phe Leu Arg Arg

180

185

190

Phe Ala Pro Ile Leu Arg Glu Val His Ala Ala Leu Ala Ala Lys Glu

195

200

205

Ala Gly Leu Arg Leu Glu Glu Thr Ala Gly Glu Gly Val Leu Arg Ala

210

215

220

Leu Glu Ala Ser Arg Ala Glu Glu Arg Glu Arg Gly Gln Thr Leu Val

Gly Pro His Arg Asp Asp Leu Val Phe Leu Leu Glu Gly Arg Pro Ala

His Arg Phe Ala Ser Arg Gly Glu Ala Lys Thr Leu Ala Leu Ala Leu

Arg Leu Ala Glu His Arg Leu Leu Gly Glu His His Gly Glu Pro Pro

Leu Leu Leu Val Asp Glu Trp Gly Glu Glu Leu Asp Glu Ala Arg Arg

Arg Ala Val Leu Ala Tyr Ala Gln Ala Leu Pro Gln Ala Ile Leu Ala

Gly Leu Glu Ala Pro Pro Gly Val Pro Val Cys Ser Val Val Arg Gly

Val Val Leu Cys Pro Gly Ala

<210> 7

<211> 2934

<212> DNA

<213> *Thermus thermophilus*

 $\langle 220 \rangle$

<221> CDS

<222> (1)..(2934)

<400> 7

atg gaa atc gcg cta gag agg atc tac ggc cac cgc ctg gcg ctc ccg 48

Met Glu Ile Ala Leu Glu Arg Ile Tyr Gly His Arg Leu Ala Leu Pro

1 5 10 15

cag gtg ggg gcg gcc ttg ctt ttc gcc cag gag gcc ccc ccg gcc ctc 96

Gln Val Gly Ala Ala Leu Leu Phe Ala Gln Glu Ala Pro Pro Ala Leu

20 25 30

ctc ctc gtc ccc gag gcg cgg ctt agg cgc tac cgg gac ctc tcc gcc 144

Leu Leu Val Pro Glu Ala Arg Leu Arg Arg Tyr Arg Asp Leu Ser Ala

35 40 45

ttc ggg gcc aag gtc tac gtg aac ccc ggc ctc gag gcc ctg gag gaa 192

Phe Gly Ala Lys Val Tyr Val Asn Pro Gly Leu Glu Ala Leu Glu Glu

50 55 60

aaa gcc ctc ttc gtc ctc tcc tac gag gag gcc cta agc ccc ttc ccc 240

Lys Ala Leu Phe Val Leu Ser Tyr Glu Glu Ala Leu Ser Pro Phe Pro

65 70 75 80

gag gac cct gag gcc tgg cgg ctt ctt ctg gag gtg ggc cgc gcc tac 288

Glu Asp Pro Glu Ala Trp Arg Leu Leu Leu Glu Val Gly Arg Ala Tyr

85 90 95

ccc cgg gag gcc ctc ctc tcc cgc ctc ctc aag ctg ggc tac gcc cgg 336

Pro Arg Glu Ala Leu Leu Ser Arg Leu Leu Lys Leu Gly Tyr Ala Arg

100

105

110

gac gag gac tac cgc gtc ctg ggg gag gtg gtg gag ctc ggc gag gtg 384

Asp Glu Asp Tyr Arg Val Leu Gly Glu Val Val Glu Leu Gly Glu Val

115

120

125

cgc ctg gag ttc ttc ggg gac gag ctg gaa agg ctt gtg gtc cgg ggg 432

Arg Leu Glu Phe Phe Gly Asp Glu Leu Glu Arg Leu Val Val Arg Gly

130

135

140

gag gaa agg cgg cgc cac gtc ctt ctg ccc aag ccg ggg aag gcg gag 480

Glu Glu Arg Arg Arg His Val Leu Leu Pro Lys Pro Gly Lys Ala Glu

145

150

155

160

ggc ttc acc tcc aag aag gtc ctc cac ttc cct ggc ccc gtc tac ctg 528

Gly Phe Thr Ser Lys Lys Val Leu His Phe Pro Gly Pro Val Tyr Leu

165

170

175

gac acc ccc gcc ctc gcc ccc aag gcc ctt tgg ccc ctc ctc gcg gga 576

Asp Thr Pro Ala Leu Ala Pro Lys Ala Leu Trp Pro Leu Leu Ala Gly

180

185

190

agg ccc tgg gtg gcc ctg ggc ggc ggg gtg gag ctc ccc ccc ttg gag 624

Arg Pro Trp Val Ala Leu Gly Gly Gly Val Glu Leu Pro Pro Leu Glu

195

200

205

ctc ggg gcg agg ccc ctt cct cct tac cgg gga agc ctg aag gcc ctg 672

Leu Gly Ala Arg Pro Leu Pro Pro Tyr Arg Gly Ser Leu Lys Ala Leu

210	215	220	
gaa aag gac ctc gcc cgc tgg ctt gcc gag ggg aag cgg gtc cac ctc 720			
Glu Lys Asp Leu Ala Arg Trp Leu Ala Glu Gly Lys Arg Val His Leu			
225	230	235	240
ttc gtg ggc cac gcc cgc acc ttg gag tac ctc aaa agg cgc ctc cag 768			
Phe Val Gly His Ala Arg Thr Leu Glu Tyr Leu Lys Arg Arg Leu Gln			
	245	250	255
gcc ttc tcg ccc ctc atc ctg gac cgc ttc ccc ggc ccc aag ggg cgg 816			
Ala Phe Ser Pro Leu Ile Leu Asp Arg Phe Pro Gly Pro Lys Gly Arg			
	260	265	270
ctt gcc ctc ctc ccc ggg gac ttt gag ggc ggg gcg gag tgg gga gag 864			
Leu Ala Leu Leu Pro Gly Asp Phe Glu Gly Gly Ala Glu Trp Gly Glu			
	275	280	285
tgg gtc ctc ctc acc gag gcc ctg gtc ttc gcc acc ggg ggg gtg cgg 912			
Trp Val Leu Leu Thr Glu Ala Leu Val Phe Ala Thr Gly Gly Val Arg			
290	295	300	
gcc agg gtc cgg gta ggg gag ggg ctc agc gac ccc ggg gcc ctt tcc 960			
Ala Arg Val Arg Val Gly Glu Gly Leu Ser Asp Pro Gly Ala Leu Ser			
305	310	315	320
cca ggg gac tac ctc atc cac ccg gag cac ggc gtc ggg cag tac ctg 1008			
Pro Gly Asp Tyr Leu Ile His Pro Glu His Gly Val Gly Gln Tyr Leu			
	325	330	335

ggc ctc gag acc cgg gag gtc ctg ggg gtc aag cgg gac tac ctg gtc 1056

Gly Leu Glu Thr Arg Glu Val Leu Gly Val Lys Arg Asp Tyr Leu Val

340

345

350

ctg cgc tac aag ggg gaa ggg aag ctc tac ctc ccc gtg gag cag ctt 1104

Leu Arg Tyr Lys Gly Glu Gly Lys Leu Tyr Leu Pro Val Glu Gln Leu

355

360

365

ccc ctc ctc aag cgc cac ccc ggg acc acc gac gac ccc ccg gag ctt 1152

Pro Leu Leu Lys Arg His Pro Gly Thr Thr Asp Asp Pro Pro Glu Leu

370

375

380

tcc tcc ctg ggc aag aac gag tgg caa agg gcc aag gag cgg gcg cgg 1200

Ser Ser Leu Gly Lys Asn Glu Trp Gln Arg Ala Lys Glu Arg Ala Arg

385

390

395

400

aag gac gtg gag gag ctg gct ggg cgc ctc ctc gtc ctc cag gcc aag 1248

Lys Asp Val Glu Glu Leu Ala Gly Arg Leu Leu Val Leu Gln Ala Lys

405

410

415

cgc aag gcc acc ccg ggc cgg gcc ttt ccc cct ttg ccc gag tgg gat 1296

Arg Lys Ala Thr Pro Gly Arg Ala Phe Pro Pro Leu Pro Glu Trp Asp

420

425

430

cct ctg gtg gag aag ggg ttc ccc tac gag ctc acc ccc gac cag aag 1344

Pro Leu Val Glu Lys Gly Phe Pro Tyr Glu Leu Thr Pro Asp Gln Lys

435

440

445

cgg gcc ctg gag gag gtc ctc cgc gac ctg gaa agc ccc cac ccc atg 1392

Arg Ala Leu Glu Glu Val Leu Arg Asp Leu Glu Ser Pro His Pro Met

450

455

460

gac cgc ctg gtc tcg ggg gac gtg ggc ttc ggc aag acg gag gtg gcc 1440

Asp Arg Leu Val Ser Gly Asp Val Gly Phe Gly Lys Thr Glu Val Ala

465

470

475

480

ctg agg gcc gcc cac cgg gtg gtg ggg cac ggg gcc cag gtg gcc ttc 1488

Leu Arg Ala Ala His Arg Val Val Gly His Gly Ala Gln Val Ala Phe

485

490

495

ctg ggg cca acc acc ctc ctc gcc gag cag cac ggg aag acc ttt agg 1536

Leu Gly Pro Thr Thr Leu Leu Ala Glu Gln His Gly Lys Thr Phe Arg

500

505

510

gag cgc ttc cag ggg ctt ccc gtg agg gtt gcg gtc ctc tcc cgc ttc 1584

Glu Arg Phe Gln Gly Leu Pro Val Arg Val Ala Val Leu Ser Arg Phe

515

520

525

acc ccg ccc aag gag gag gag gcc atc cta aaa ggc ctc gcc gag ggg 1632

Thr Pro Pro Lys Glu Glu Glu Ala Ile Leu Lys Gly Leu Ala Glu Gly

530

535

540

acg gtg gac atc gtc atc ggc acc cac cgc ctc ctc cag gag gac gtg 1680

Thr Val Asp Ile Val Ile Gly Thr His Arg Leu Leu Gln Glu Asp Val

545

550

555

560

cgc ttc agg gac ctc ggc ctc ctc atc gtg gac gag gag cac cgc ttc 1728

Arg Phe Arg Asp Leu Gly Leu Leu Ile Val Asp Glu Glu His Arg Phe

565

570

575

ggc gtg gcc caa aag gag agg atc cgg gag ctc aag gcg gag gtg gac 1776

Gly Val Ala Gln Lys Glu Arg Ile Arg Glu Leu Lys Ala Glu Val Asp

580

585

590

acc ctc tac ctc tcc gcc acc ccc atc ccc cgc acc ctc tac tcc gcc 1824

Thr Leu Tyr Leu Ser Ala Thr Pro Ile Pro Arg Thr Leu Tyr Ser Ala

595

600

605

ctg gtg ggc ctc aaa gac ctt tcc agc atc cag acc ccg ccc ccg ggg 1872

Leu Val Gly Leu Lys Asp Leu Ser Ser Ile Gln Thr Pro Pro Pro Gly

610

615

620

cgc aag ccc atc aag acc ttc ctc gct ccc ttt gat ccc ctc ttg gtg 1920

Arg Lys Pro Ile Lys Thr Phe Leu Ala Pro Phe Asp Pro Leu Leu Val

625

630

635

640

cgg gag gcc atc ctc ttt gag ctg gag cgt ggg ggc aag gtc ttc tac 1968

Arg Glu Ala Ile Leu Phe Glu Leu Glu Arg Gly Gly Lys Val Phe Tyr

645

650

655

gtc cac gac cgg gtg gcc tcc ata gag gcc agg cgg cgc ttt ctg gaa 2016

Val His Asp Arg Val Ala Ser Ile Glu Ala Arg Arg Arg Phe Leu Glu

660

665

670

aac ctc gtc ccc gag gcc cgc atc ggg gtg gtc cac ggc cag atg ccc 2064

Asn Leu Val Pro Glu Ala Arg Ile Gly Val Val His Gly Gln Met Pro

675	680	685	
gaa agc ctc att gag gag acc atg ctc ctc ttc gcc gaa ggg gcg tac			2112
Glu Ser Leu Ile Glu Glu Thr Met Leu Leu Phe Ala Glu Gly Ala Tyr			
690	695	700	
gac gtc ctc ctc gcc acc acc atc att gag gcg ggc ctg gac gtg ccc			2160
Asp Val Leu Leu Ala Thr Thr Ile Ile Glu Ala Gly Leu Asp Val Pro			
705	710	715	720
gag gcg aac acc atc ctc att gag cgg gcg gac cgc ctg ggc ctc gcc			2208
Glu Ala Asn Thr Ile Leu Ile Glu Arg Ala Asp Arg Leu Gly Leu Ala			
725	730	735	
acc ttg tac cag ctc cgg ggc cgg gtg ggg cgg agg gag gag gag gcc			2256
Thr Leu Tyr Gln Leu Arg Gly Arg Val Gly Arg Arg Glu Glu Glu Ala			
740	745	750	
tac gcc tac ctc ttc cac ccg cct cgc ctc acc gag gcc gcg gag aag			2304
Tyr Ala Tyr Leu Phe His Pro Pro Arg Leu Thr Glu Ala Ala Glu Lys			
755	760	765	
cgc ctc gcc gcc atc gcc gac ctc tcc gat ctg ggc tcg ggc cac ctc			2352
Arg Leu Ala Ala Ile Ala Asp Leu Ser Asp Leu Gly Ser Gly His Leu			
770	775	780	
ctg gcc gaa agg gac atg gaa atc cgg ggc gtg ggg aac ctt ttg ggg			2400
Leu Ala Glu Arg Asp Met Glu Ile Arg Gly Val Gly Asn Leu Leu Gly			
785	790	795	800

ccg gag cag cac ggg cac atc cgg gcg ctt tcc ctc gag gtc tac acc 2448

Pro Glu Gln His Gly His Ile Arg Ala Leu Ser Leu Glu Val Tyr Thr

805

810

815

gag ctt ctg gaa gag gcc atc cgc aag ctc aag ggg gag gcc aag gag 2496

Glu Leu Leu Glu Glu Ala Ile Arg Lys Leu Lys Gly Glu Ala Lys Glu

820

825

830

gag cgg cgg cac gtg acc ctg gac ctc gcc ctc tcc gcc cgg ctg ccc 2544

Glu Arg Arg His Val Thr Leu Asp Leu Ala Leu Ser Ala Arg Leu Pro

835

840

845

gcg gag tac gtg ggg agc ctc gag gcc agg agc cgc tac tac agc cgt 2592

Ala Glu Tyr Val Gly Ser Leu Glu Ala Arg Ser Arg Tyr Tyr Ser Arg

850

855

860

ttt gcc gag gcg aaa agc ctc gcc gag ctt tcc cgc ctg gtg cgg gag 2640

Phe Ala Glu Ala Lys Ser Leu Ala Glu Leu Ser Arg Leu Val Arg Glu

865

870

875

880

ctc aaa gag cgc tac ggg ccc ctt cct gag gag gcg gag aac ttc gtg 2688

Leu Lys Glu Arg Tyr Gly Pro Leu Pro Glu Glu Ala Glu Asn Phe Val

885

890

895

gcc ctc gcc cgg ctc cgc ctg gtg gcg gag agg aag ggg gtg gtg tcc 2736

Ala Leu Ala Arg Leu Arg Leu Val Ala Glu Arg Lys Gly Val Val Ser

900

905

910

atc acg gag ggc ctc acc cac ctg gag gtg gtc ttc ccc cgc tac ccc 2784

Ile Thr Glu Gly Leu Thr His Leu Glu Val Val Phe Pro Arg Tyr Pro

915

920

925

ctg gac tac gac gcc cgc ggc ctc aag ggg ctt ccc tac cgg gtg gag 2832

Leu Asp Tyr Asp Ala Arg Gly Leu Lys Gly Leu Pro Tyr Arg Val Glu

930

935

940

ctt acg cag tac ccg ccc ggg ttc cgc ctg gag aag aag ggc ctg agg 2880

Leu Thr Gln Tyr Pro Pro Gly Phe Arg Leu Glu Lys Lys Gly Leu Arg

945

950

955

960

ccc cgg gac tac ccc gag gcc ctg atg gag gtg ctc tac ctc ttc gcc 2928

Pro Arg Asp Tyr Pro Glu Ala Leu Met Glu Val Leu Tyr Leu Phe Ala

965

970

975

gac ctc

2934

Asp Leu

<210> 8

<211> 978

<212> PRT

<213> *Thermus thermophilus*

<400> 8

Met Glu Ile Ala Leu Glu Arg Ile Tyr Gly His Arg Leu Ala Leu Pro

1

5

10

15

Gln Val Gly Ala Ala Leu Leu Phe Ala Gln Glu Ala Pro Pro Ala Leu

20

25

30

Leu Leu Val Pro Glu Ala Arg Leu Arg Arg Tyr Arg Asp Leu Ser Ala

35

40

45

Phe Gly Ala Lys Val Tyr Val Asn Pro Gly Leu Glu Ala Leu Glu Glu

50

55

60

Lys Ala Leu Phe Val Leu Ser Tyr Glu Glu Ala Leu Ser Pro Phe Pro

65

70

75

80

Glu Asp Pro Glu Ala Trp Arg Leu Leu Leu Glu Val Gly Arg Ala Tyr

85

90

95

Pro Arg Glu Ala Leu Leu Ser Arg Leu Leu Lys Leu Gly Tyr Ala Arg

100

105

110

Asp Glu Asp Tyr Arg Val Leu Gly Glu Val Val Glu Leu Gly Glu Val

115

120

125

Arg Leu Glu Phe Phe Gly Asp Glu Leu Glu Arg Leu Val Val Arg Gly

130

135

140

Glu Glu Arg Arg Arg His Val Leu Leu Pro Lys Pro Gly Lys Ala Glu

145

150

155

160

Gly Phe Thr Ser Lys Lys Val Leu His Phe Pro Gly Pro Val Tyr Leu

165

170

175

Asp Thr Pro Ala Leu Ala Pro Lys Ala Leu Trp Pro Leu Leu Ala Gly

180

185

190

Arg Pro Trp Val Ala Leu Gly Gly Gly Val Glu Leu Pro Pro Leu Glu

195

200

205

Leu Gly Ala Arg Pro Leu Pro Pro Tyr Arg Gly Ser Leu Lys Ala Leu

210

215

220

Glu Lys Asp Leu Ala Arg Trp Leu Ala Glu Gly Lys Arg Val His Leu

225

230

235

240

Phe Val Gly His Ala Arg Thr Leu Glu Tyr Leu Lys Arg Arg Leu Gln

245

250

255

Ala Phe Ser Pro Leu Ile Leu Asp Arg Phe Pro Gly Pro Lys Gly Arg

260

265

270

Leu Ala Leu Leu Pro Gly Asp Phe Glu Gly Gly Ala Glu Trp Gly Glu

275

280

285

Trp Val Leu Leu Thr Glu Ala Leu Val Phe Ala Thr Gly Gly Val Arg

290

295

300

Ala Arg Val Arg Val Gly Glu Gly Leu Ser Asp Pro Gly Ala Leu Ser

305

310

315

320

Pro Gly Asp Tyr Leu Ile His Pro Glu His Gly Val Gly Gln Tyr Leu

325

330

335

Gly Leu Glu Thr Arg Glu Val Leu Gly Val Lys Arg Asp Tyr Leu Val

340

345

350

Leu Arg Tyr Lys Gly Glu Gly Lys Leu Tyr Leu Pro Val Glu Gln Leu

355

360

365

Pro Leu Leu Lys Arg His Pro Gly Thr Thr Asp Asp Pro Pro Glu Leu

370

375

380

Ser Ser Leu Gly Lys Asn Glu Trp Gln Arg Ala Lys Glu Arg Ala Arg

385

390

395

400

Lys Asp Val Glu Glu Leu Ala Gly Arg Leu Leu Val Leu Gln Ala Lys

405

410

415

Arg Lys Ala Thr Pro Gly Arg Ala Phe Pro Pro Leu Pro Glu Trp Asp

420

425

430

Pro Leu Val Glu Lys Gly Phe Pro Tyr Glu Leu Thr Pro Asp Gln Lys

435

440

445

Arg Ala Leu Glu Glu Val Leu Arg Asp Leu Glu Ser Pro His Pro Met

450

455

460

Asp Arg Leu Val Ser Gly Asp Val Gly Phe Gly Lys Thr Glu Val Ala

465

470

475

480

Leu Arg Ala Ala His Arg Val Val Gly His Gly Ala Gln Val Ala Phe

485

490

495

Leu Gly Pro Thr Thr Leu Leu Ala Glu Gln His Gly Lys Thr Phe Arg

500

505

510

Glu Arg Phe Gln Gly Leu Pro Val Arg Val Ala Val Leu Ser Arg Phe

515

520

525

Thr Pro Pro Lys Glu Glu Glu Ala Ile Leu Lys Gly Leu Ala Glu Gly

530

535

540

Thr Val Asp Ile Val Ile Gly Thr His Arg Leu Leu Gln Glu Asp Val

545

550

555

560

Arg Phe Arg Asp Leu Gly Leu Leu Ile Val Asp Glu Glu His Arg Phe

565

570

575

Gly Val Ala Gln Lys Glu Arg Ile Arg Glu Leu Lys Ala Glu Val Asp

580

585

590

Thr Leu Tyr Leu Ser Ala Thr Pro Ile Pro Arg Thr Leu Tyr Ser Ala

595

600

605

Leu Val Gly Leu Lys Asp Leu Ser Ser Ile Gln Thr Pro Pro Pro Gly

610

615

620

Arg Lys Pro Ile Lys Thr Phe Leu Ala Pro Phe Asp Pro Leu Leu Val

625

630

635

640

Arg Glu Ala Ile Leu Phe Glu Leu Glu Arg Gly Gly Lys Val Phe Tyr

645

650

655

Val His Asp Arg Val Ala Ser Ile Glu Ala Arg Arg Arg Phe Leu Glu

660

665

670

Asn Leu Val Pro Glu Ala Arg Ile Gly Val Val His Gly Gln Met Pro

675

680

685

Glu Ser Leu Ile Glu Glu Thr Met Leu Leu Phe Ala Glu Gly Ala Tyr

690

695

700

Asp Val Leu Leu Ala Thr Thr Ile Ile Glu Ala Gly Leu Asp Val Pro

705

710

715

720

Glu Ala Asn Thr Ile Leu Ile Glu Arg Ala Asp Arg Leu Gly Leu Ala

725

730

735

Thr Leu Tyr Gln Leu Arg Gly Arg Val Gly Arg Arg Glu Glu Glu Ala

740

745

750

Tyr Ala Tyr Leu Phe His Pro Pro Arg Leu Thr Glu Ala Ala Glu Lys

755

760

765

Arg Leu Ala Ala Ile Ala Asp Leu Ser Asp Leu Gly Ser Gly His Leu

770

775

780

Leu Ala Glu Arg Asp Met Glu Ile Arg Gly Val Gly Asn Leu Leu Gly

785

790

795

800

Pro Glu Gln His Gly His Ile Arg Ala Leu Ser Leu Glu Val Tyr Thr

805

810

815

Glu Leu Leu Glu Glu Ala Ile Arg Lys Leu Lys Gly Glu Ala Lys Glu

820

825

830

Glu Arg Arg His Val Thr Leu Asp Leu Ala Leu Ser Ala Arg Leu Pro

835

840

845

Ala Glu Tyr Val Gly Ser Leu Glu Ala Arg Ser Arg Tyr Tyr Ser Arg

850

855

860

Phe Ala Glu Ala Lys Ser Leu Ala Glu Leu Ser Arg Leu Val Arg Glu

865

870

875

880

Leu Lys Glu Arg Tyr Gly Pro Leu Pro Glu Glu Ala Glu Asn Phe Val

885

890

895

Ala Leu Ala Arg Leu Arg Leu Val Ala Glu Arg Lys Gly Val Val Ser

900

905

910

Ile Thr Glu Gly Leu Thr His Leu Glu Val Val Phe Pro Arg Tyr Pro

915

920

925

Leu Asp Tyr Asp Ala Arg Gly Leu Lys Gly Leu Pro Tyr Arg Val Glu

930

935

940

Leu Thr Gln Tyr Pro Pro Gly Phe Arg Leu Glu Lys Lys Gly Leu Arg

945

950

955

960

Pro Arg Asp Tyr Pro Glu Ala Leu Met Glu Val Leu Tyr Leu Phe Ala

965

970

975

Asp Leu

<210> 9

<211> 38

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: synthetic DNA

<400> 9

atatcatatg gaagcctggc ggaaagccct cctgcct

38

<210> 10

<211> 41

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: synthetic DNA

<400> 10

atatagatct ttattatgcg tccgggaggg ggactacgcc c

41

<210> 11

<211> 31

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: synthetic DNA

<400> 11

atcatatgag agaccggggtc cgctggcggg t

31

<210> 12

<211> 32

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: synthetic DNA

<400> 12

atagatcttt acaggtccac cgcctggacc tc

32

<210> 13

<211> 49

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: synthetic DNA

<400> 13

actacttggt acactgacgc gagcacgcag gagctcattc cagtgcgca 49

<210> 14

<211> 38

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: synthetic DNA

<400> 14

atatcatatg cgtcttctcc tcttcggca acggaact 38

<210> 15

<211> 41

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: synthetic DNA

<400> 15

atatagatct ttattaggcg ccagggcaca ggaccacccc t 41

<210> 16

<211> 38

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: synthetic DNA

<400> 16

atatcatatg gaaatcgcg tagagaggat ctacggcc

38

<210> 17

<211> 41

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: synthetic DNA

<400> 17

atatagatct ttattagagg tcggcgaaga ggtagagcac c

41

【 0 1 1 7 】

【配列表フリーテキスト】

配列番号 9 : 合成DNA

配列番号10 : 合成DNA

配列番号11 : 合成DNA

配列番号12 : 合成DNA

配列番号13 : 合成DNA

配列番号14 : 合成DNA

配列番号15 : 合成DNA

配列番号16：合成DNA

配列番号17：合成DNA

【図面の簡単な説明】

【図 1】

MutYの機能を示す図である。

【図 2】

MutYの塩基除去修復機構を示す図である。

【図 3】

MutYのSDS-ポリアクリルアミドゲル電気泳動の結果を示す写真である。

【図 4】

MutYのゲルろ過の結果を示す図である。

【図 5】

MutYのアミノ酸配列の比較結果を示す図である。

【図 6】

MutYの活性の測定法の概要を示す図である。

【図 7】

MutYの基質特異性を示す図である。

【図 8】

MutYの吸収スペクトルを示す図である。

【図 9】

MutYのCDスペクトルを示す図である。

【図 1 0】

MutYの熱安定性を示す図である。

【図 1 1】

基質DNA及び³²P標識部位を示す図である。

【図 1 2】

RecJの機能を示す図である。

【図 1 3】

RecJのSDS-ポリアクリルアミドゲル電気泳動の結果を示す写真である。

【図 1 4】

RecJのアミノ酸配列を比較した図である。

【図 1 5】

RecJのCDスペクトルを示す図である。

【図 1 6】

RecJの熱安定性を示す図である。

【図 1 7】

RecJのエキソヌクレアーゼ活性の測定法を示す図である。

【図 1 8】

RecJのエキソヌクレアーゼ活性の結果を示す図である。

【図 1 9】

RecJのエキソヌクレアーゼ活性の結果を示す図である（RecJ濃度依存性）。

【図 2 0】

エテノヌクレオチドのRecJの活性に与える影響を示す図である。

【図 2 1】

RecJのエキソヌクレアーゼ活性の結果を示す図である（蛍光スペクトル）。

【図 2 2】

RecJのエキソヌクレアーゼ活性の結果を示す図である（蛍光強度及び蛍光偏光度の時間変化）。

【図 2 3】

RecJのエキソヌクレアーゼ活性の結果を示す図である（DNA濃度依存性）。

【図 2 4】

RecFの反応経路を示す図である。

【図 2 5】

RecFのSDS-ポリアクリルアミドゲル電気泳動の結果を示す写真である。

【図 2 6】

RecFのゲルろ過を行った結果を示す写真である。

【図 2 7】

RecFのアミノ酸配列を比較した図である。

【図 2 8】

RecFと ϵ DNAとの結合を示す図である。

【図 2 9】

ATPase活性を示す図である。

【図 3 0】

ATPase活性を示す図である（DNA依存性）。

【図 3 1】

TRCFのヌクレオチド除去修復機構を示す図である。

【図 3 2】

UvrBの立体構造を示す図である。

【図 3 3】

TRCF- β 及びUvrB- β のSDS-ポリアクリルアミドゲル電気泳動の結果を示す写真である。

【図 3 4】

TRCF- β 及びUvrB- β のアミノ酸配列の比較を示す図である。

【図 3 5】

TRCF- β 及びUvrB- β のSDスペクトルを示す図である。

【図 3 6】

TRCF- β 及びUvrB- β の熱安定性を示す図である。

【図 3 7】

TRCF- β 及びUvrB- β のpH安定性を示す図である。

【図 3 8】

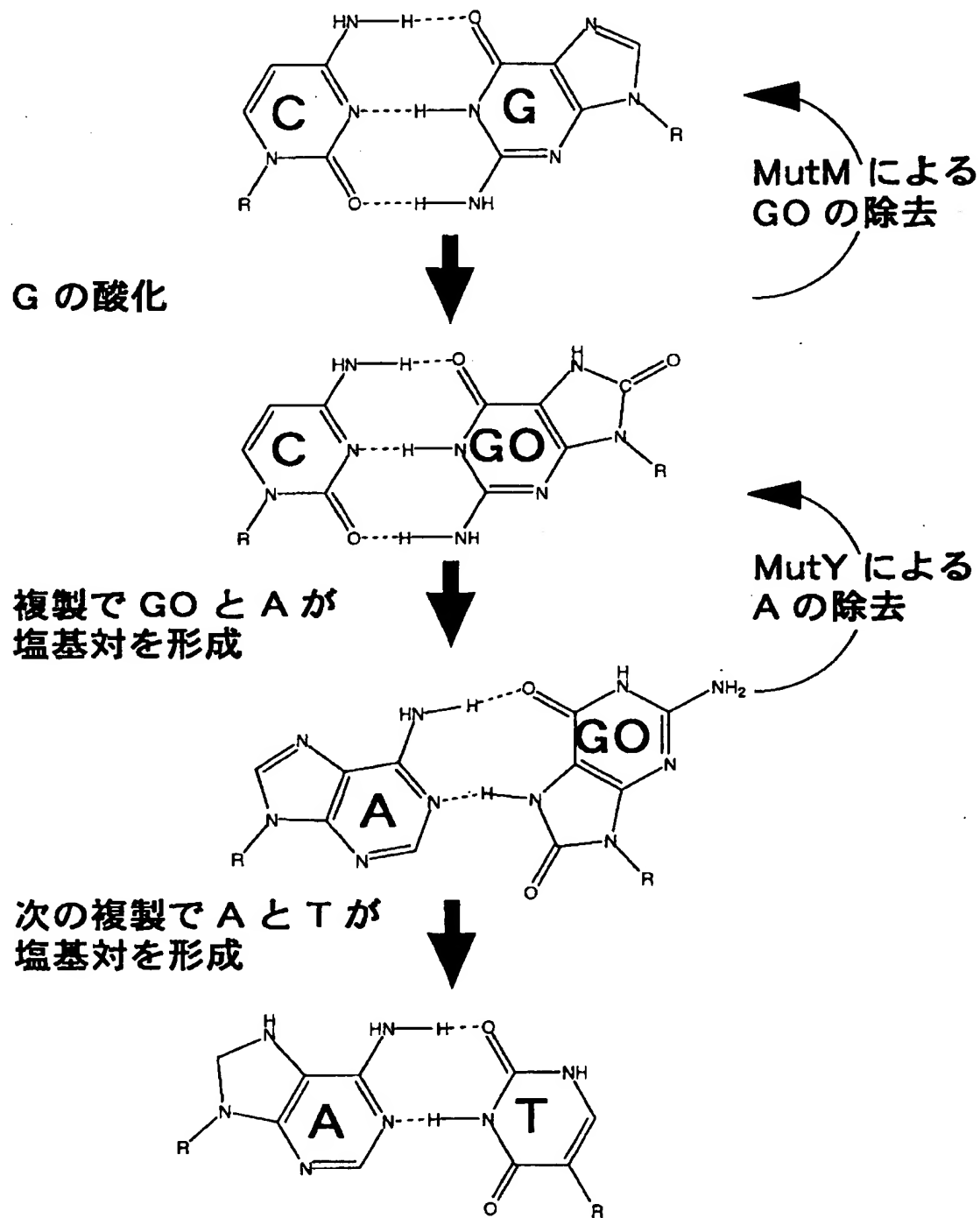
BIA CoreによりTRCFの相互作用を測定した結果を示す図である。

【図 3 9】

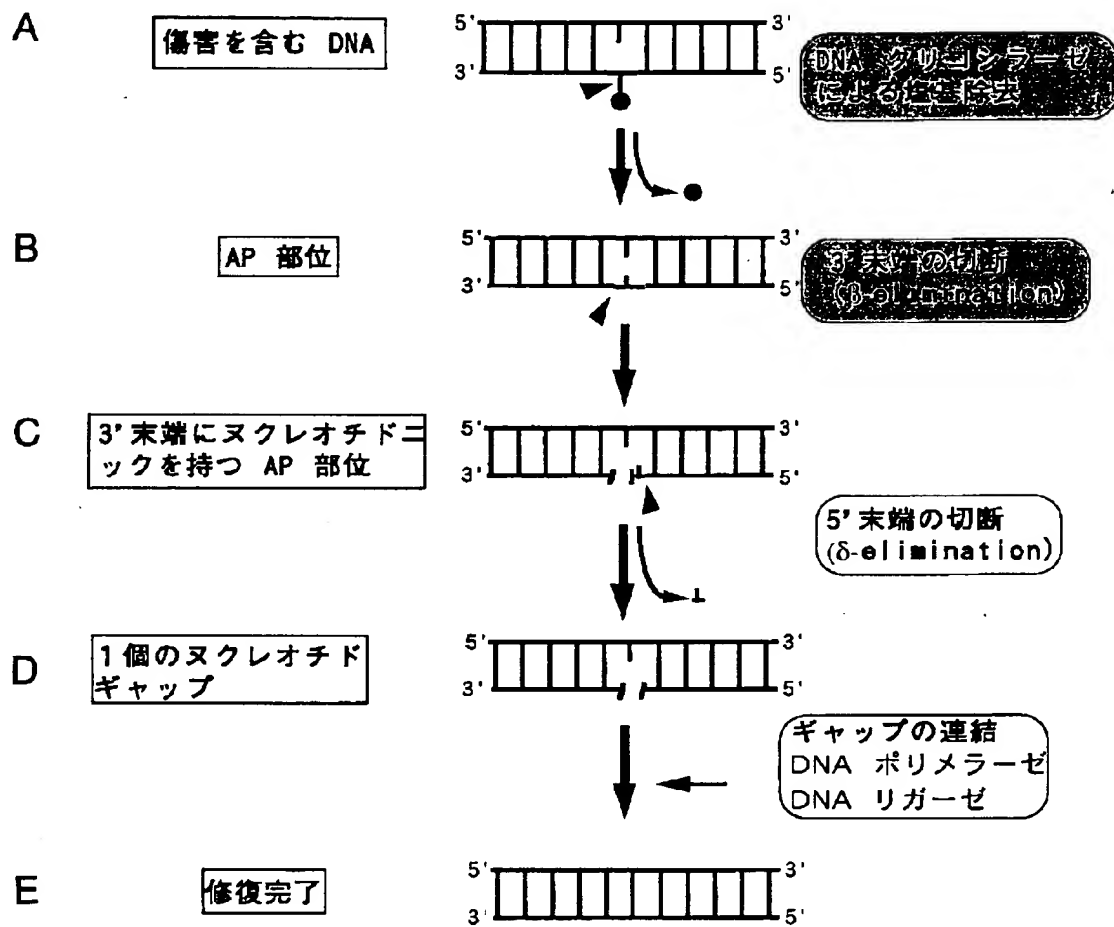
TRCFとUvrAとの相互作用を測定した結果を示す図である。

【書類名】 図面

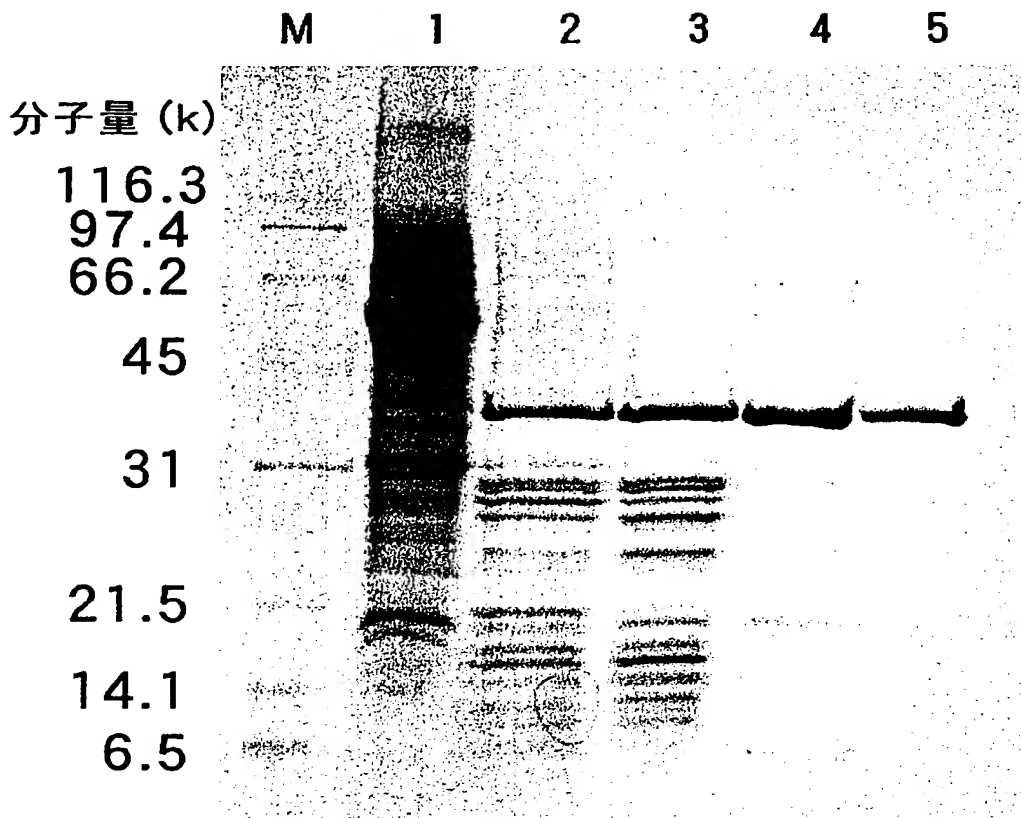
【図1】



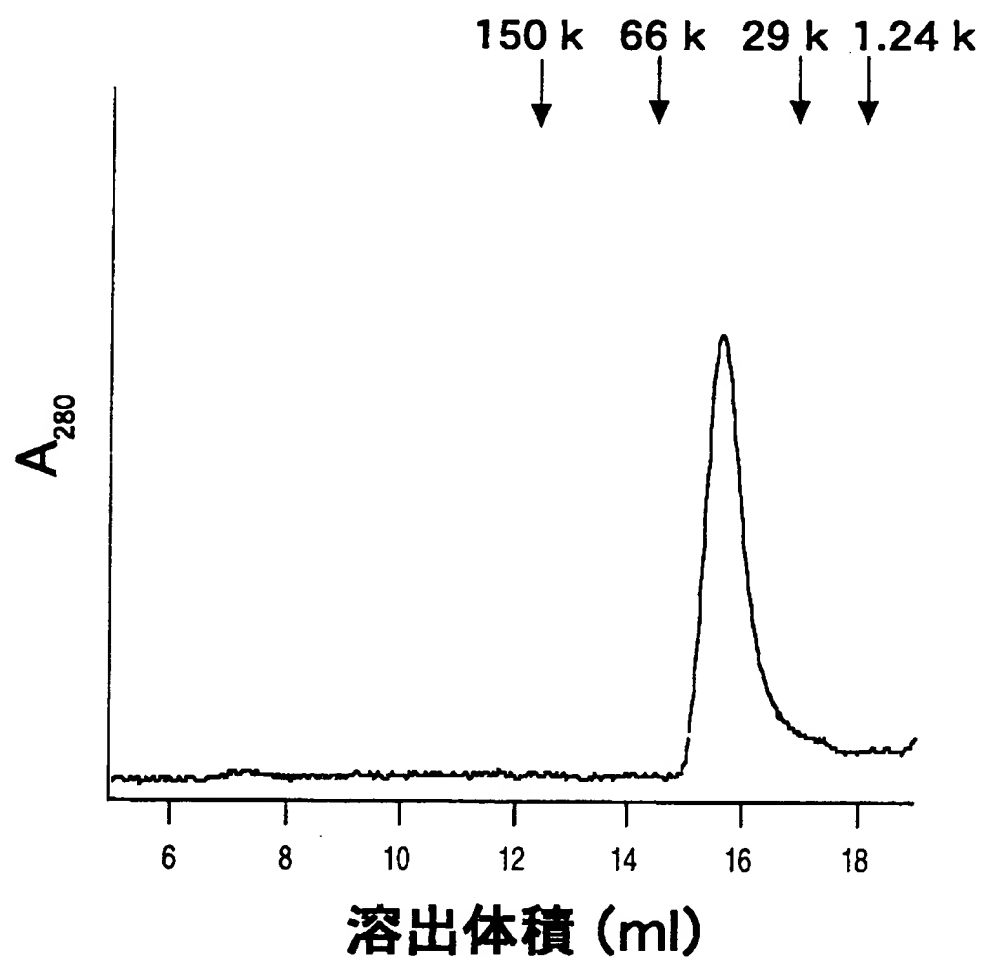
【図 2】



【図 3】



【図 4】



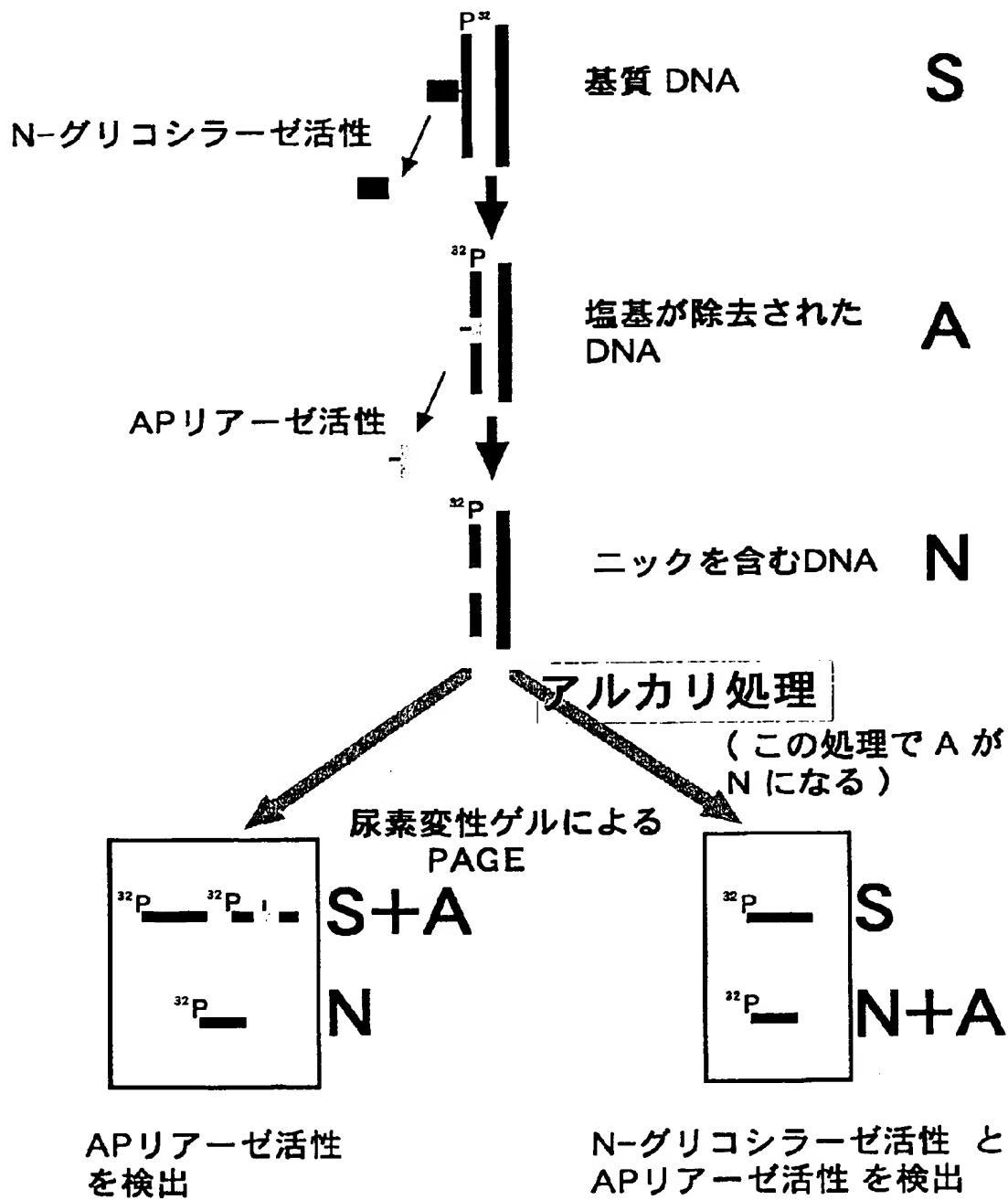
【図 5】

Tth MutY	1	MEAMKALANWYEN-APPWPR-GE-KDPYRVLVSEVLLQOTRVEQALPYRRFL	53
Hsa MutY	51	CDGLARQPEEVLQASVSSVHLFRDAEVAETAFRCSSLSWDOE-KRDLPHRRRADEMDLD	139
Spo MutY	1	MSDSNHFLDLHSYTOLEVERFRESLHOFDXT-KRILPHRRKECIPSEDSPLDEWEPVORLYEVLVSEIMLQOTRVTVKRYTKNM	88
Eco MutY	1	MDASOFSAQVNDWQYGRKTLPMO-ID-KTPYKVMSEVMLQOTQVATVIPYFERFM	56
Eco EndoIII	1	MNKAKRLIEITLREN-NPHPTT-ELN-FSSPFELLIAVLISAOATDVSNKATAKLY	55
Tth MutY	54	ERFRITUKALKAASLE-EVLRWQAGYYR-RAEHLHRLARSVEEL-PPSFAELR-GLPGLPYTAAVAASHAEGEVAAVGNVRRVLSRLFARES	145
Hsa MutY	140	OKMPTLQDLASASLE-EVNLWAGLGYYS-FGRRLQEGKPKVVEELGGHPRTAETQOLPGVGRYTAGAIAIAFGQATGMGNVAVLCRVPAIGA	237
Spo MutY	89	ETLPTIKSCAEENYNTQVMPVNSGFTY-ICKRLHQAQGLAKLHPSEIPRTGDENWAGIPGVGPYTAGAVLSIAWKOPTGIVGNVIRVLSRALAIHS	187
Eco MutY	57	ARFPTVTLNAPID-EVLHLMTGLGYA-FARNLHKAQOVATLHGKFPETFEVA-ALPGVGRSTAGATLSLKGHPFILLGNVKNVLCARCYAVSG	153
Eco EndoIII	56	PVANTPAAMLELGE-GVKTYIKTIGLYNSKAENIKTGRILLEHNGEVPEDRAALE-ALPGVGRKTANVNLNTAFQWPTIAVTHIFRYCNRTQFAPG	153
Tth MutY	146	-PK-EKELEALQAQGLPEGVDPGVWNOALNIEIGATVCLKPRPGACPLGAFCRG-KEAPGRYP-APR	210
Hsa MutY	238	DPSSTLVSOOLWGLAQOQVDP-ARPGDENDAAIELGATVCTPORPLCSQCPVESLCRARORVEOELLASGSLSGSPDVEECAPNTGCOCHLCLPPSEPMD	336
Spo MutY	188	DCSKGKANALINKLANELVDP-VRPGDFQVALNIEIGATICTPQSPRCVCPISEICKAYQ-EQNVIRDGMTKYD-IEDVPCN-ICITDIPS-K	276
Eco MutY	154	WPGKKEVENKLNSEQVTPA-VGVERFQAMDLGAMICTRSKPKCSLQPLQNGCIA-AANNSWALYP-GKK-P-K	225
Eco EndoIII	154	-KN-VEQVEEKLKVP-A-EFKVDCHHMLJLHGRYTCIARKPRGSCIJEDLCEY-*	205
Tth MutY	211	RRAK-EER-LVALVLLGRKG-VHLREGR-FQGLYGVLPFP-ELP-GREAFGVRS-RP	266
Hsa MutY	337	QTLGV-VNFPKASRKPPREESATCVLEOPGA-LGAQILLYVORNSGLAGMEEPSTVM-PSSEOLQRKALLQELORWAGP-LPATHLRHL	425
Spo MutY	277	EDLQNWARYPVHPAKTKORE-ERALWIFOKTDPSTKEKFFLIRKRBSAGILAGWIDEPTIEFGQESWPKMDIDAEFOKSIQOMISNDSRSLIKKYOSR	375
Eco MutY	226	QTL-PER-TGYFELLQH-EDEVLAQHPPSGLWGLYCFQFAD-EES-LRQMLAORO-IADNLTQL	287
Eco EndoIII	206	-K-EKVD	
Tth MutY	267	GEVRHALTHRRLR-VEVR-GALIEGEGEDPWKRP-LPKLMELRLKALP-PLAH-AGWPLPDA	325
Hsa MutY	426	GEVHTFESHIKLYOVYGLALEGQTPVTTPPGARMLTQEEHTAAVSTAMKVFVWQGOQPGTGMGSKRSQVSSPCSRKKPRMGQOVLNDFRSHISTDAHSLNSAAQ	535
Spo MutY	376	GRYLHIFSHIRKTSHFVYAIAS-PDVTNEDFFWISQSLFHVGNIC-ELGLKNVRAALEIKRK-VTSLN-FKEPKLTSARRIVTKAEC461	
Eco MutY	288	TAFRHTFSHFHLD-IYP-MMLPVSSFTGCMDEGNAWYNLQAP-PSVG-LAAPVER-LLOOLRTGAPV	350

Tth (*Thermus thermophilus* HB8), Hsa (*Homo sapiens*), Spo (*Schizosaccharomyces pombe*), Eco (*Escherichia coli*)

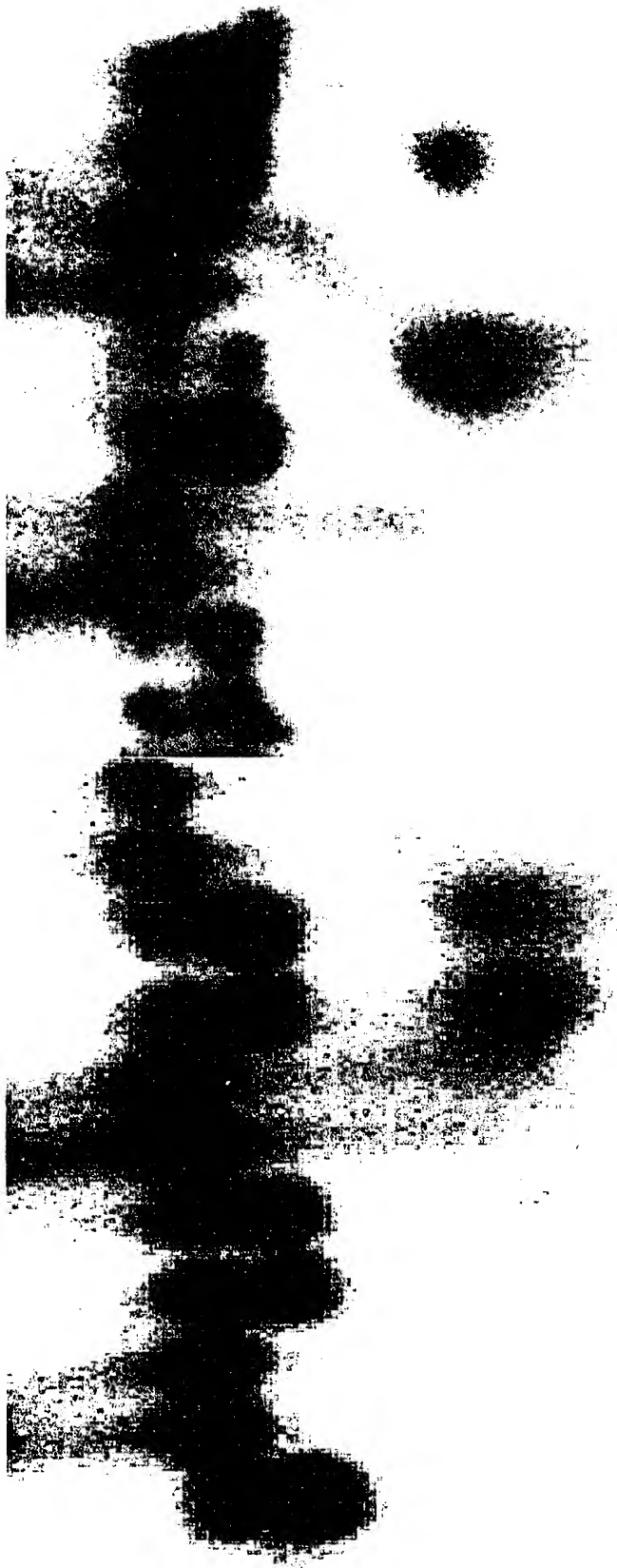
N-グリコシラーゼ活性に必須な残基 * 鉄硫黄クラスターを構成する残基 (D)

【図 6】

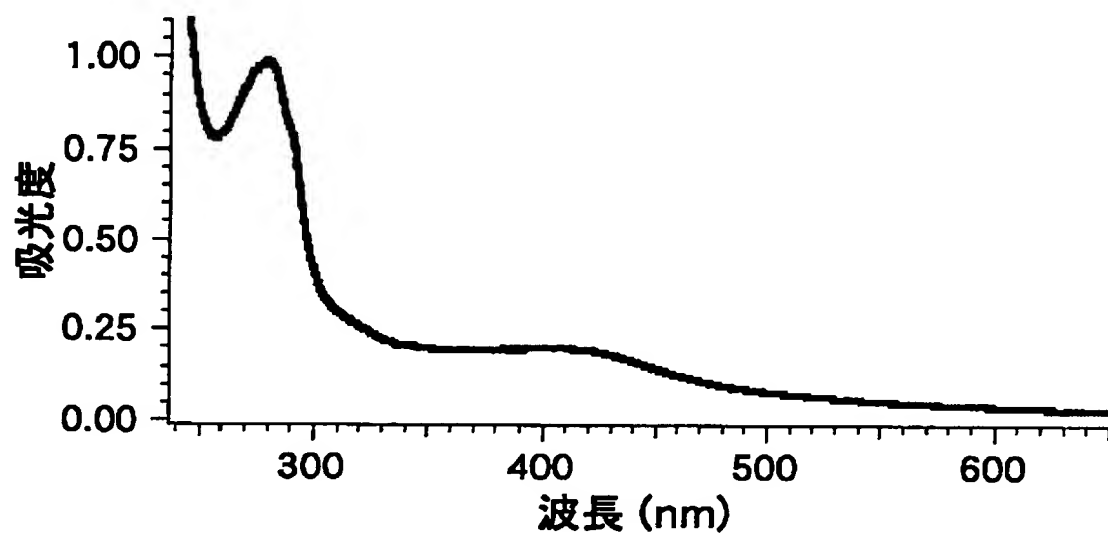


【図 7】

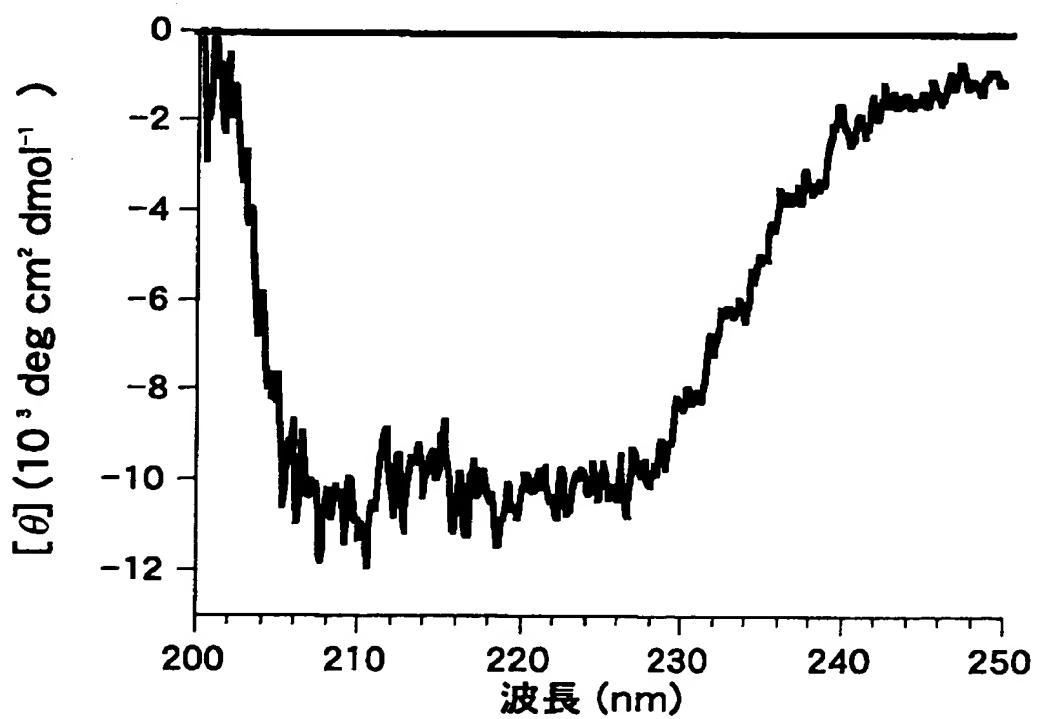
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21



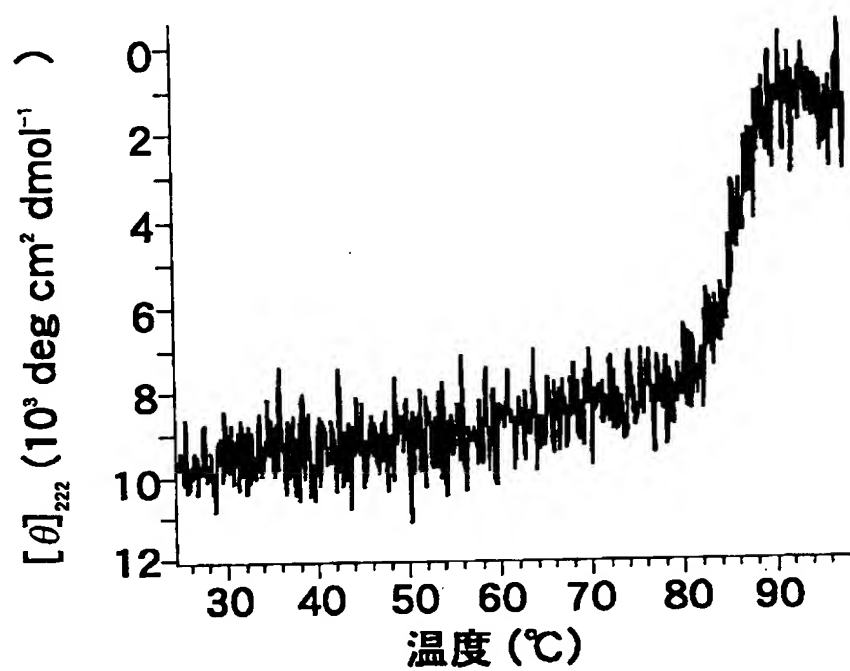
【図 8】



【図 9】



【図 1 0】



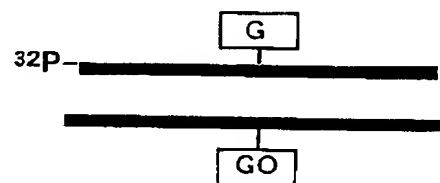
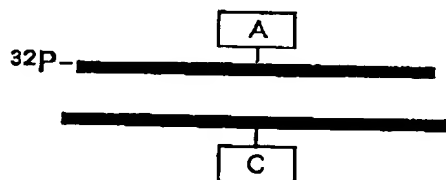
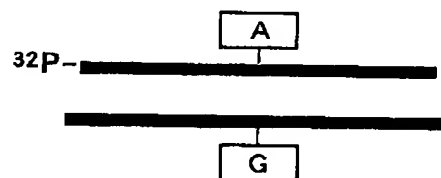
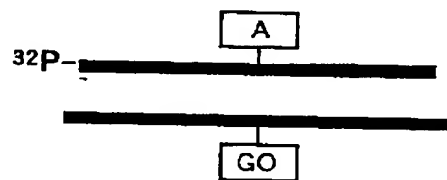
【図 1 1】

5' - [32P] AGATCTTGACGGGGAAAYCCGAATTCGGCGAACGTGGCGAG - 3'
 3' - AATCTAGAACTGCCCCCTTTXGGCTTAAGCCGCTTGCACCGCTCTT - 5'

X : GO, G, C, T

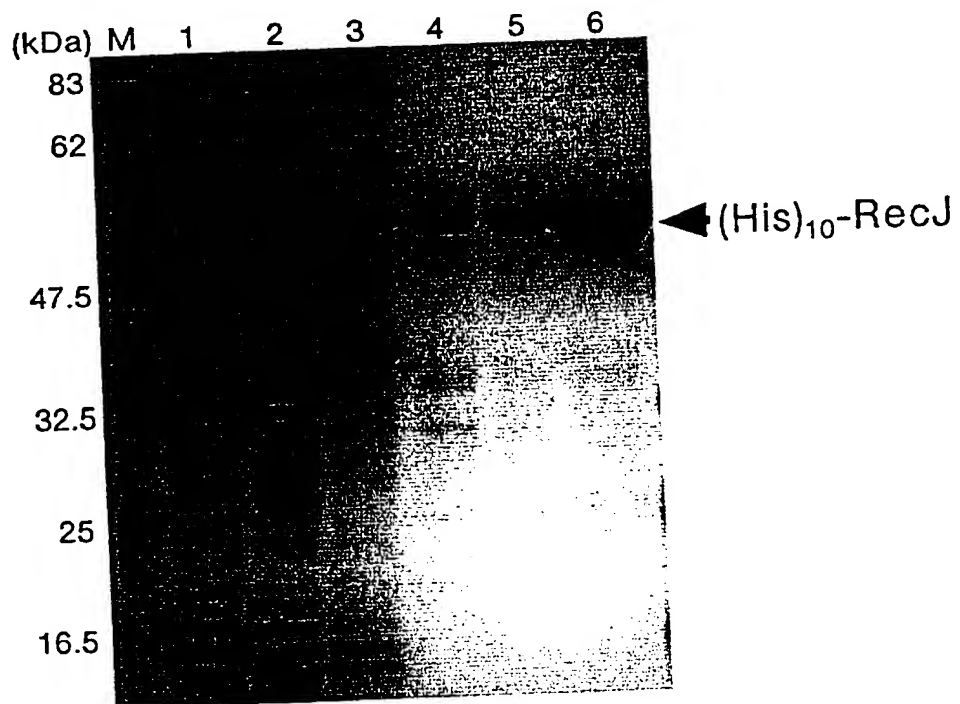
Y : A, G

↓
アニーリング



【図 1 2】

【図 13】



【図 14】

Motif I

RecJ_Tt	[73]	KRIRVHG	DY	A	G	L	T	G	T	A	I	L	V	R	G	L	A	A	L	G	[100]	
RecJ_Ec	[73]	TRIIVVG	D	F	A	G	A	T	S	T	A	L	S	V	L	A	M	R	S	L	G	[100]
RecJ_Aa	[56]	KRIIIYG	D	Y	V	G	I	T	G	T	A	I	L	Y	R	V	L	K	L	L	G	[79]
RecJ_Hp	[47]	TEILVVG	D	Y	A	G	V	I	S	S	A	I	M	A	K	F	F	E	S	L	N	[74]
RecJ_Hi	[67]	QKIVIVG	D	F	A	G	A	T	S	T	A	L	S	V	L	A	L	R	Q	L	G	[90]

PPX1_Sc	[29]	TICVGNES	A	M	S	I	A	S	A	I	T	S	Y	C	Q	Y	I	Y	N	[52]
PRUNE_Dm	[37]	HLVMGNES	C	L	S	A	V	S	A	V	T	L	A	F	V	Y	A	A	S	[60]

Motif II

RecJ_Tt	[128]	SDLFLTV	C	G	I	T	N	H	A	E	L	R	E	[147]
RecJ_Ec	[131]	AQLIVTV	N	G	I	S	S	H	A	G	V	E	H	[150]
RecJ_Aa	[133]	GDFLITV	N	G	T	S	A	V	E	E	I	D	Q	[152]
RecJ_Hp	[102]	APLITTV	N	G	I	N	A	F	E	A	A	R	F	[121]
RecJ_Hi	[126]	VQLLMTV	N	G	V	S	S	F	D	G	V	A	F	[145]

PPX1_Sc	[120]	ELNSYLV	N	N	D	T	P	K	N	L	K	N	Y	[139]
PRUNE_Dm	[87]	PLVCEMV	C	R	A	R	V	A	L	P	R	R	Y	[106]

Motif III

[153]	VEVIVT	T	P	G	K	[165]	
[155]	IPVIVT	P	L	P	G	D	[165]
[154]	LETVVI	N	V	P	P	[164]	
[126]	YTLIIT	C	L	H	H	[136]	
[150]	IRVLVT	L	P	P	E	[151]	

[141]	NVVGII	F	D	L	Q	[153]	
[128]	NVTEIL	R	P	L	E	D	[140]

Motif IV

RecJ_Tt	[209]	YADLA	A	V	G	T	I	A	V	A	P	L	W	G	W	[228]	
RecJ_Ec	[226]	LLDL	V	A	L	G	T	V	A	V	V	P	L	D	A	N	[245]
RecJ_Aa	[215]	FLDL	V	A	L	G	L	L	A	Y	M	P	V	N	P	V	[234]
RecJ_Hp	[189]	LLCL	A	G	V	A	T	I	A	M	M	P	L	T	F	F	[208]
RecJ_Hi	[219]	LLDL	V	A	L	G	T	I	A	V	V	P	L	D	Q	N	[238]

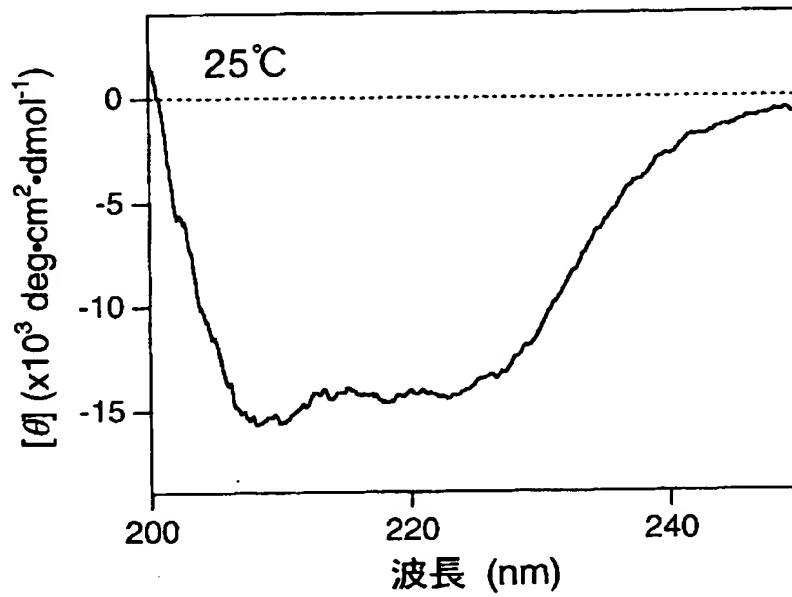
PPX1_Sc	[191]	IALL	L	M	G	A	I	I	T	S	N	M	R	R	K	[210]	
PRUNE_Dm	[183]	VAQL	L	H	A	T	I	V	L	T	I	N	F	A	P	A	[202]

Specific motif

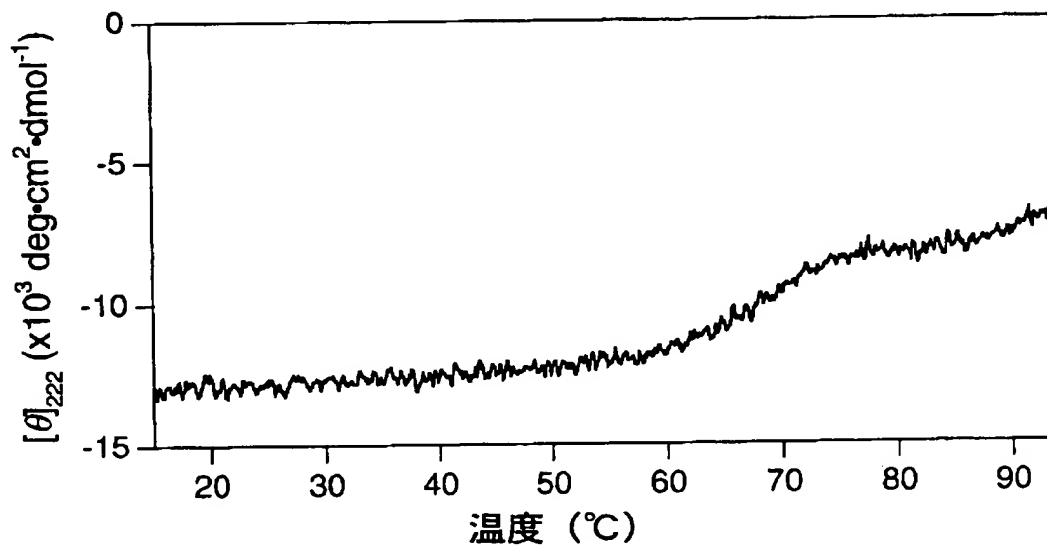
[386]	DLLLR	Y	K	E	A	A	G	F	A	M	[402]
[421]	GMMLK	F	A	M	A	A	G	L	S	L	[438]
[404]	DMFLK	W	D	K	A	M	G	L	T	L	[420]
[372]	SLLLG	Y	R	Q	A	C	G	L	S	V	[388]
[415]	NMILK	F	A	M	A	A	G	L	S	I	[431]

Tt : *Thermus thermophilus* HB8, Ec : *Escherichia coli*, Aa : *Aquifex aeolicus*,
 Hp : *Helicobacter pylori*, Hi : *Haemophilus influenzae* Rd,
 Sc : *Saccharomyces cerevisiae*, Dm : *Drosophila melanogaster*

【図 15】



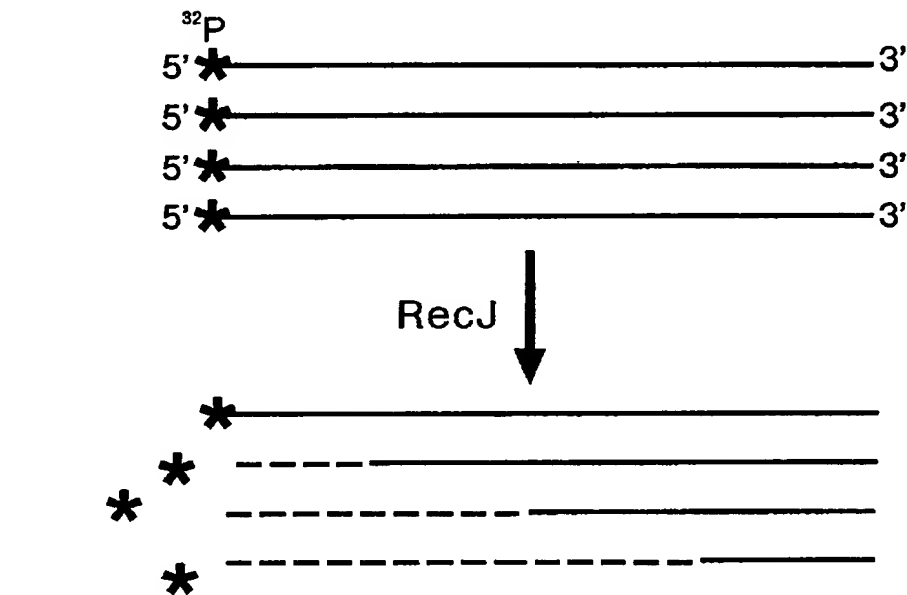
【図 16】



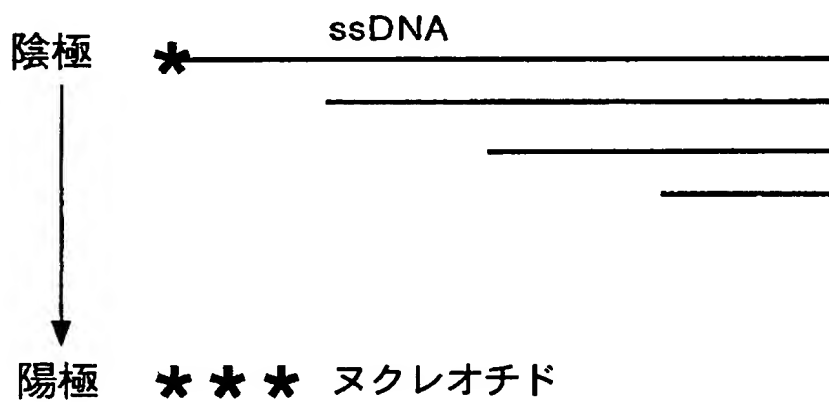
【図 1 7】

基質 DNA : 49-mer ssDNA

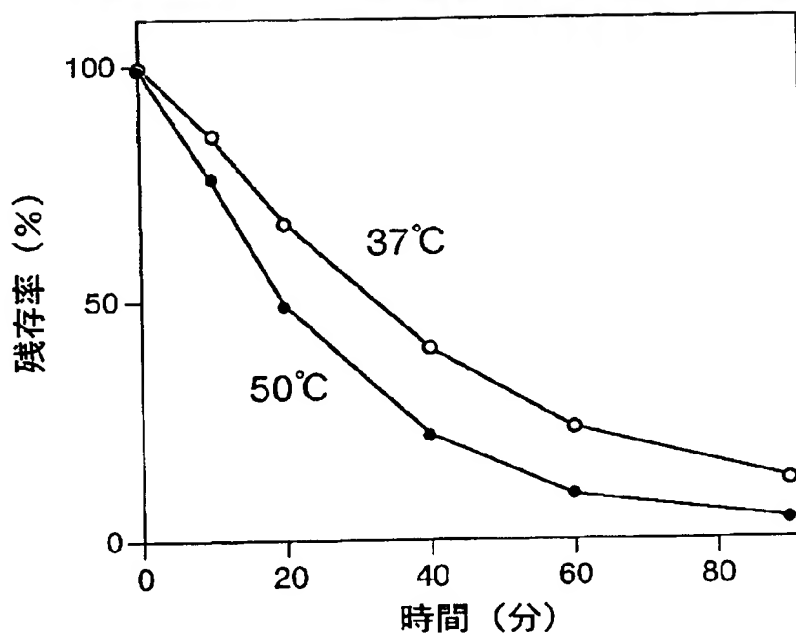
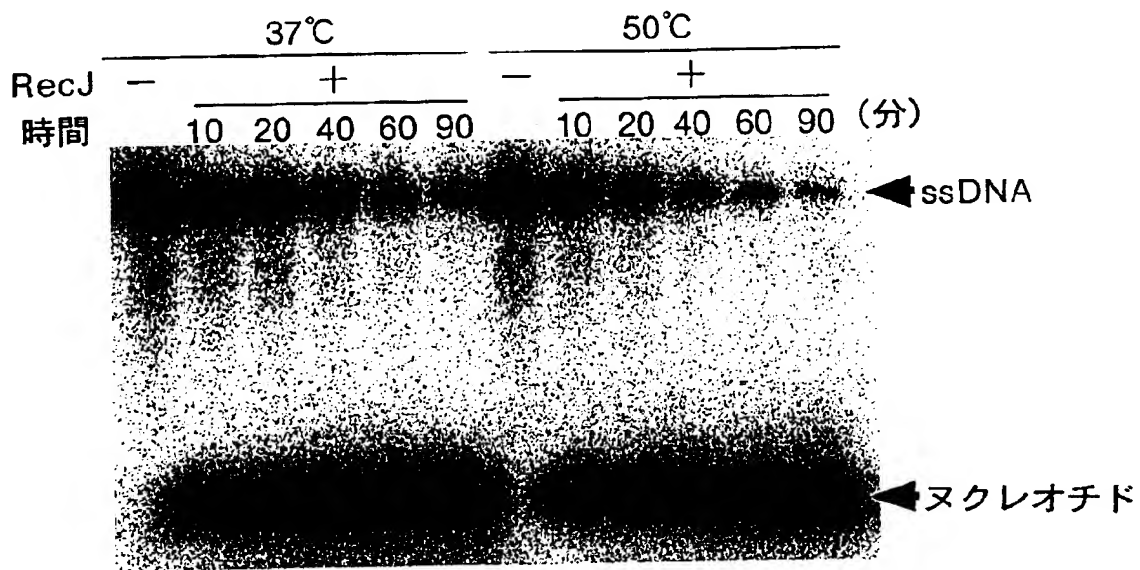
5'-ACTACTTGGTACACTGACGCGAGCACGCAGGAGCTCATTCCAGTGCGCA-3'



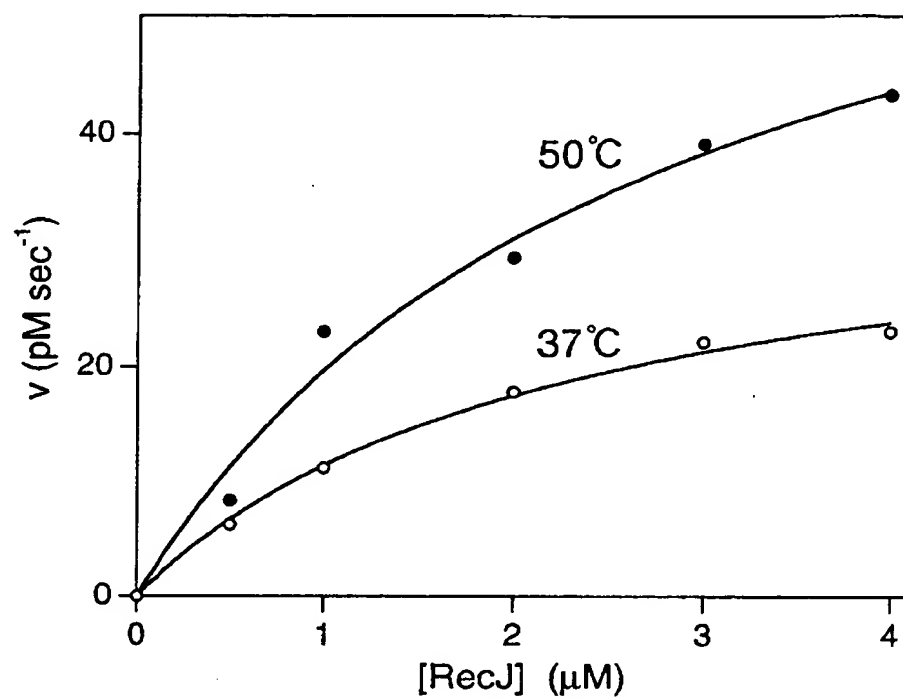
7 M Urea を含む変性ゲル
でのゲル電気泳動



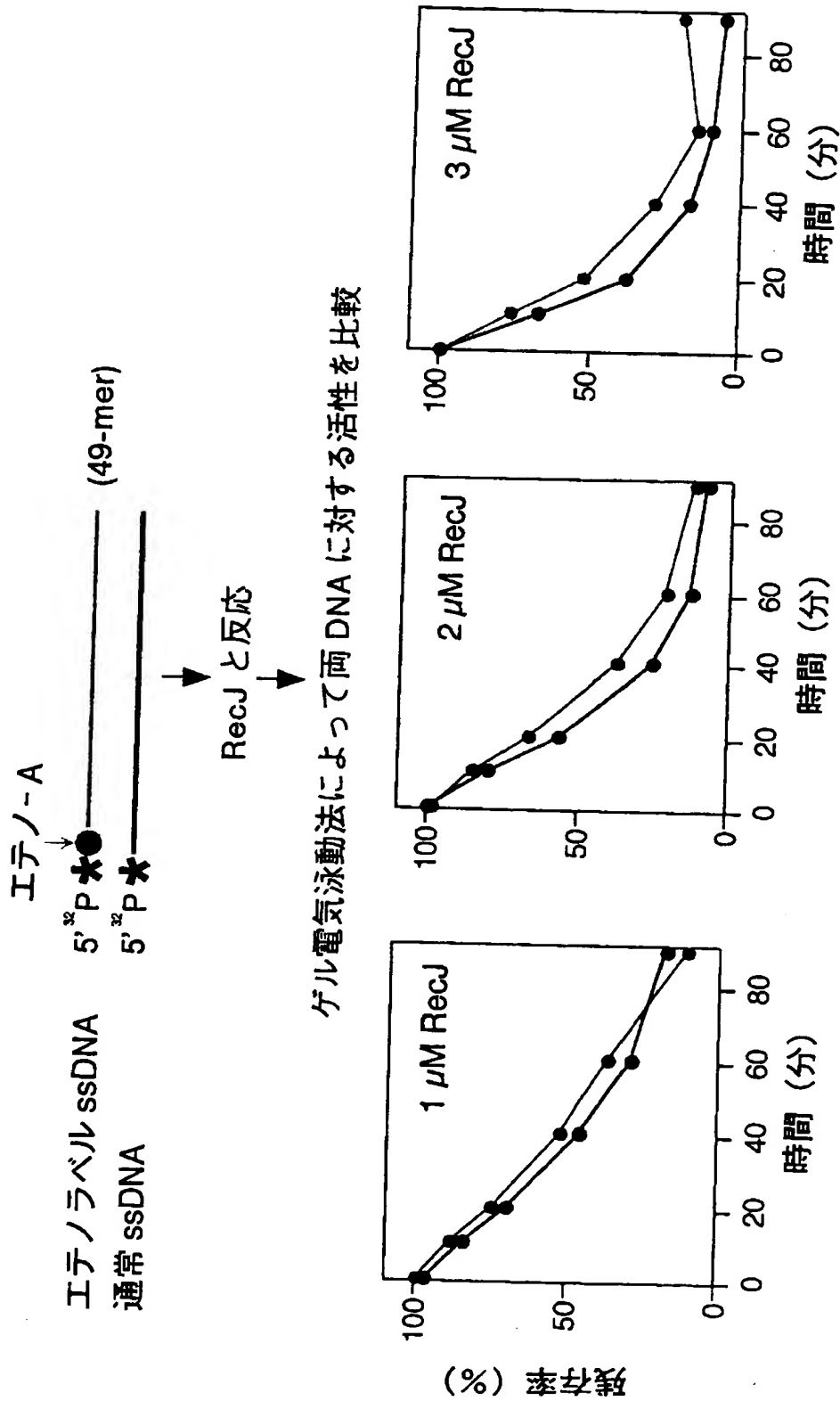
【図18】



【図19】

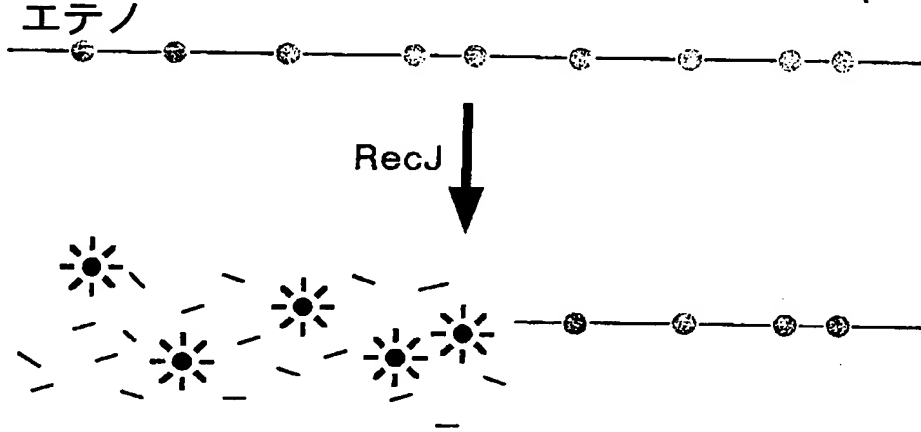


【図 20】

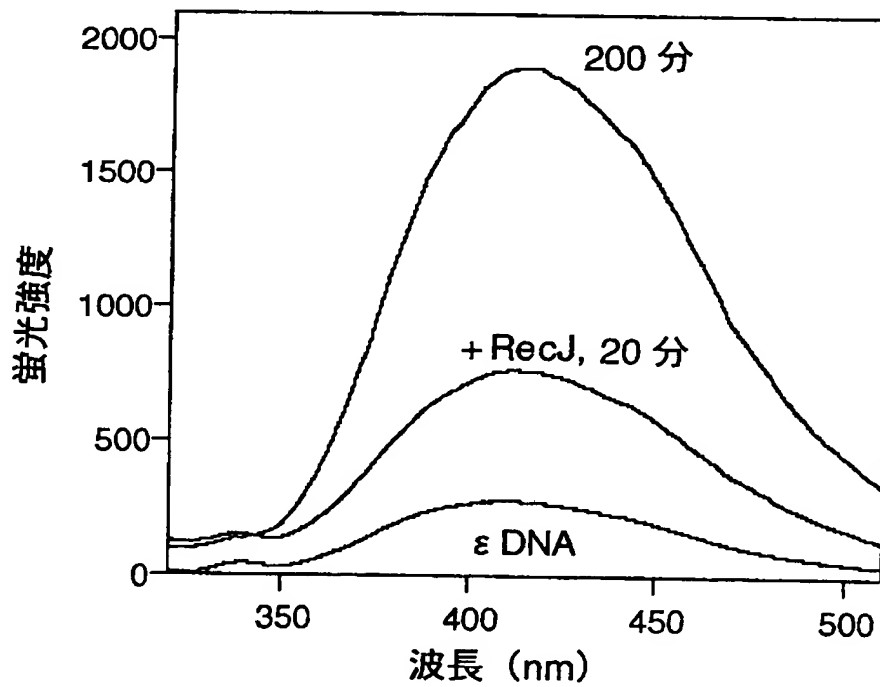


【図 21】

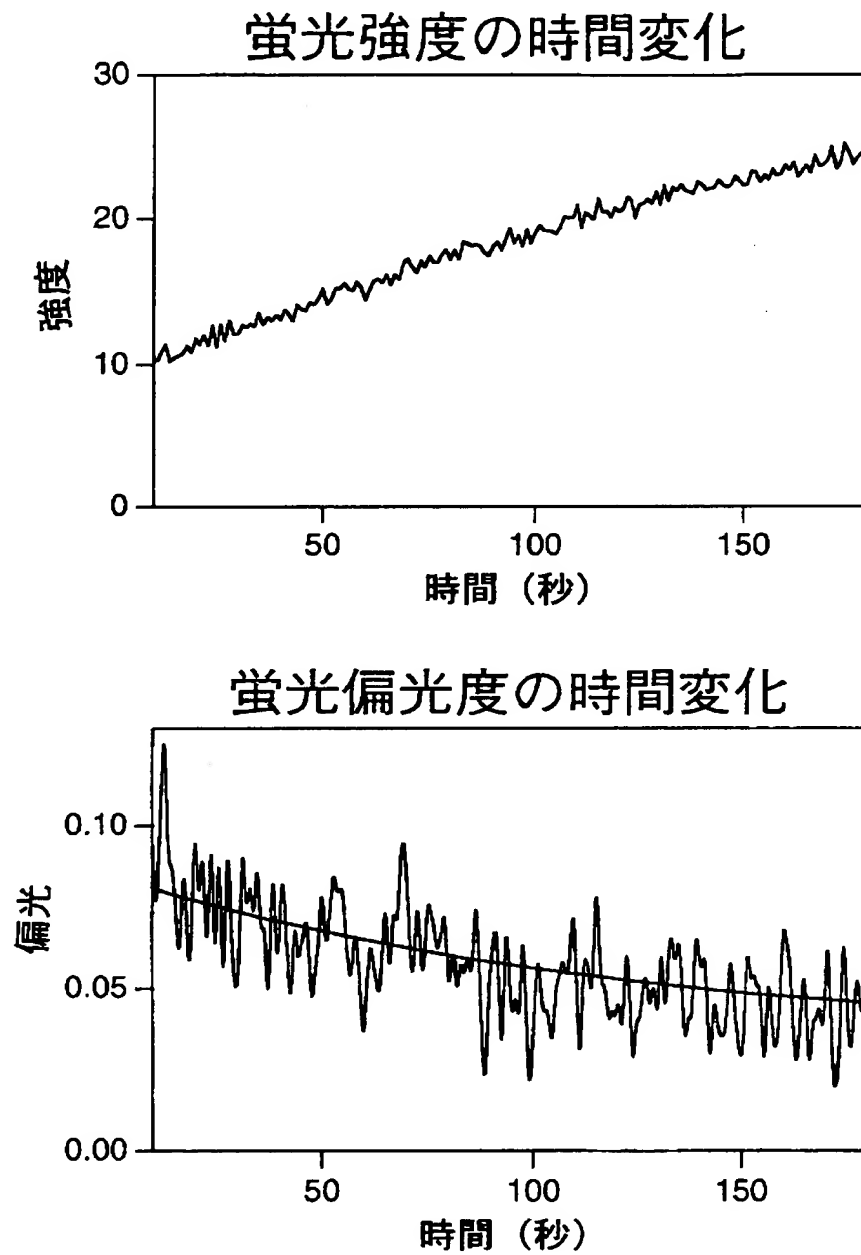
基質 DNA: エテノラベル ウシ胸腺 ssDNA (ϵ DNA)



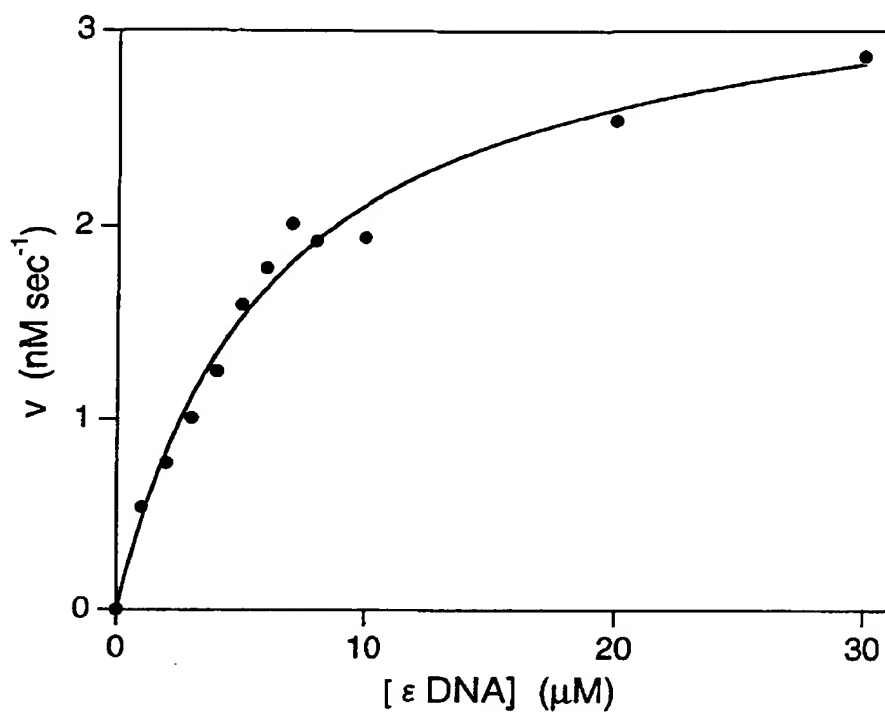
蛍光スペクトル



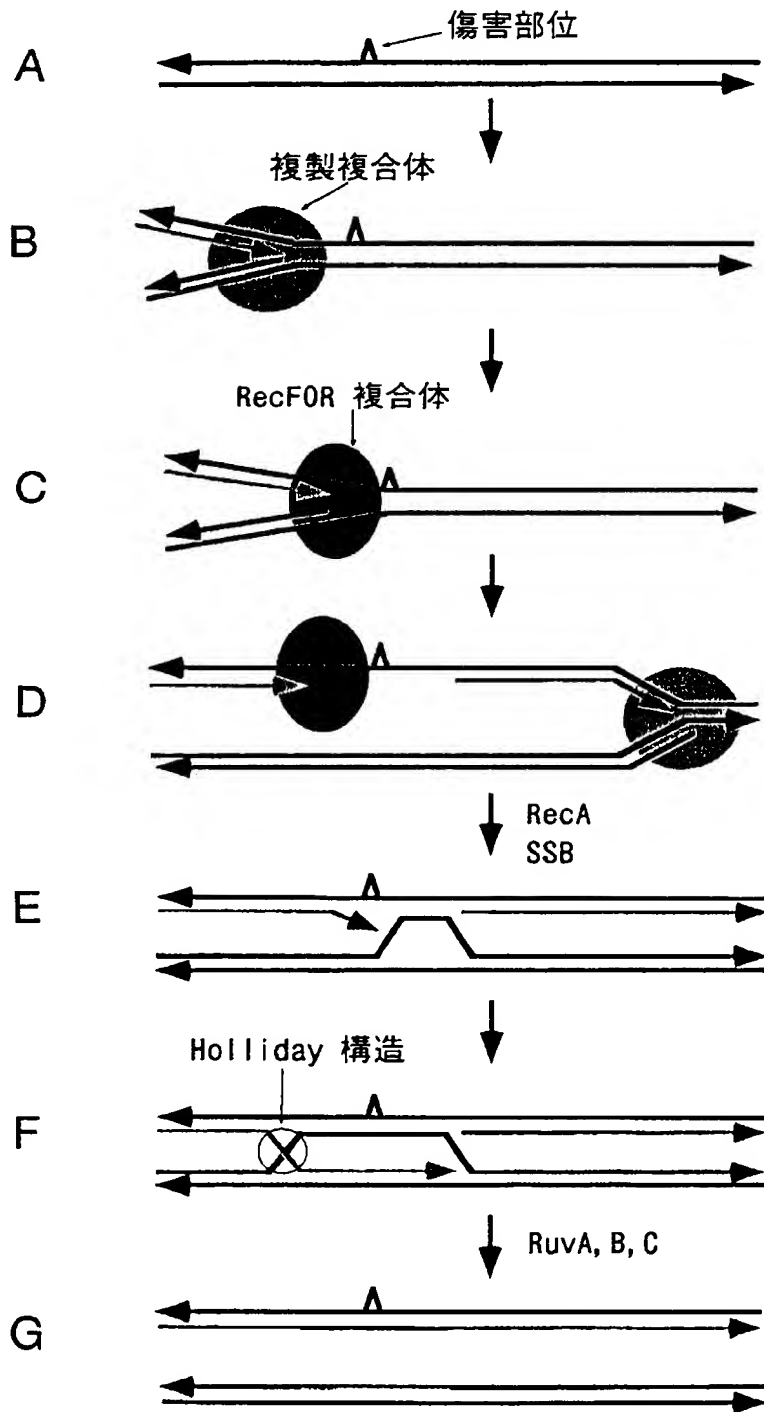
【図 2 2】



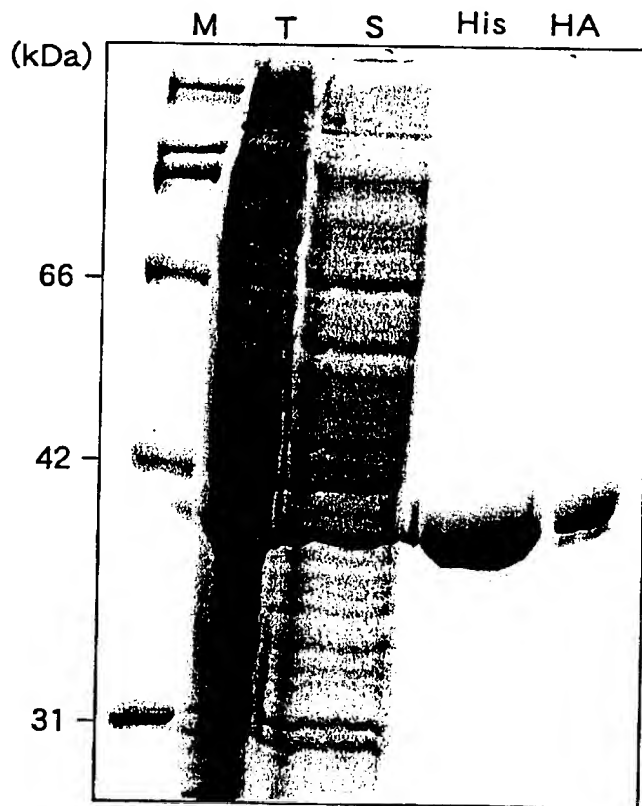
【図 23】



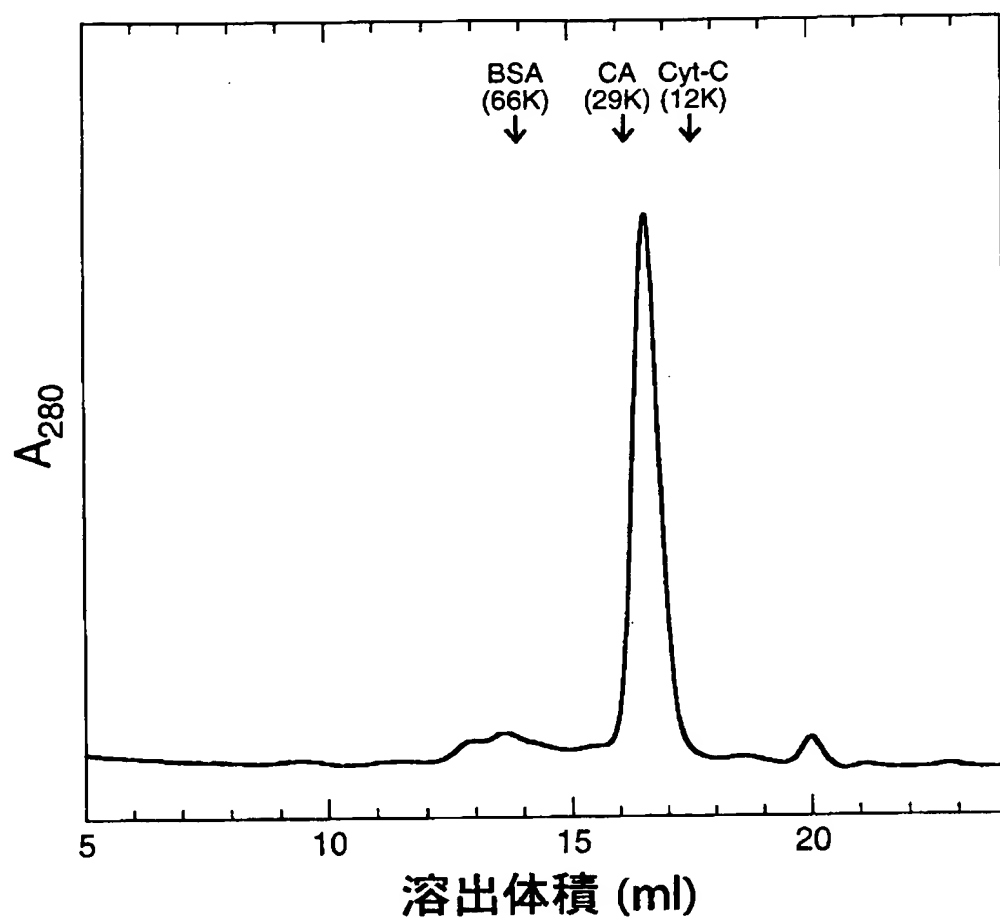
【図 24】



【図 25】



【図 2 6】

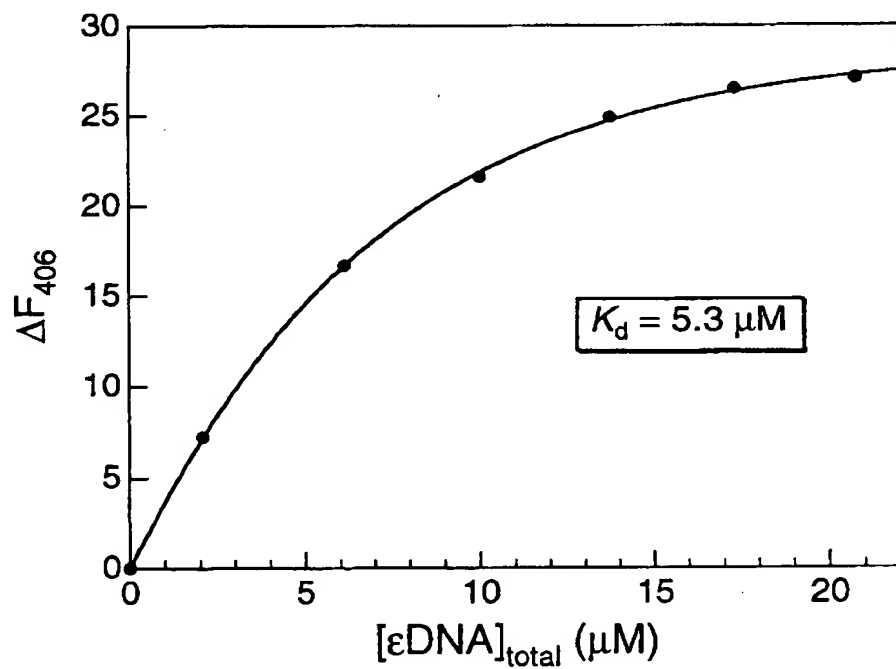
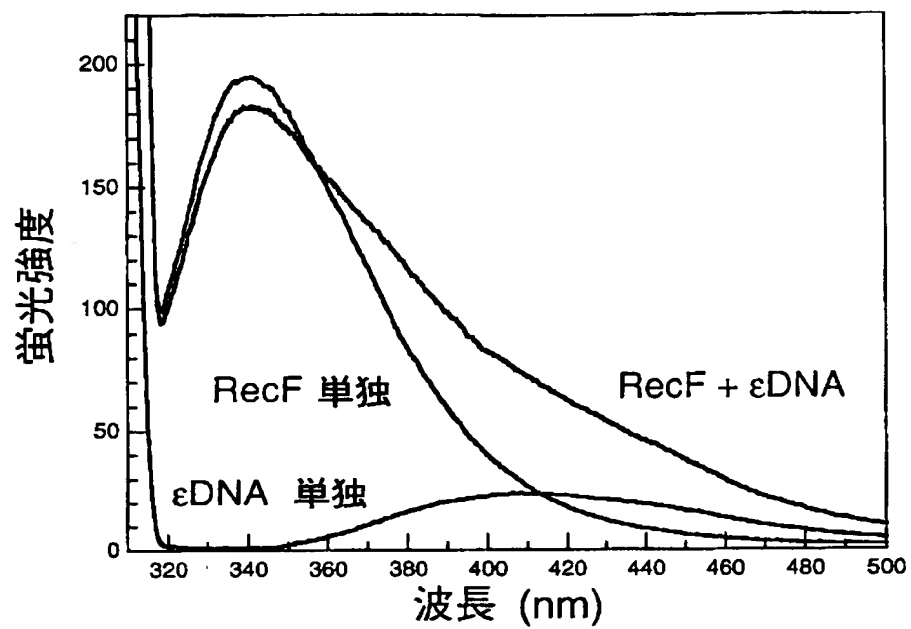


【図27】

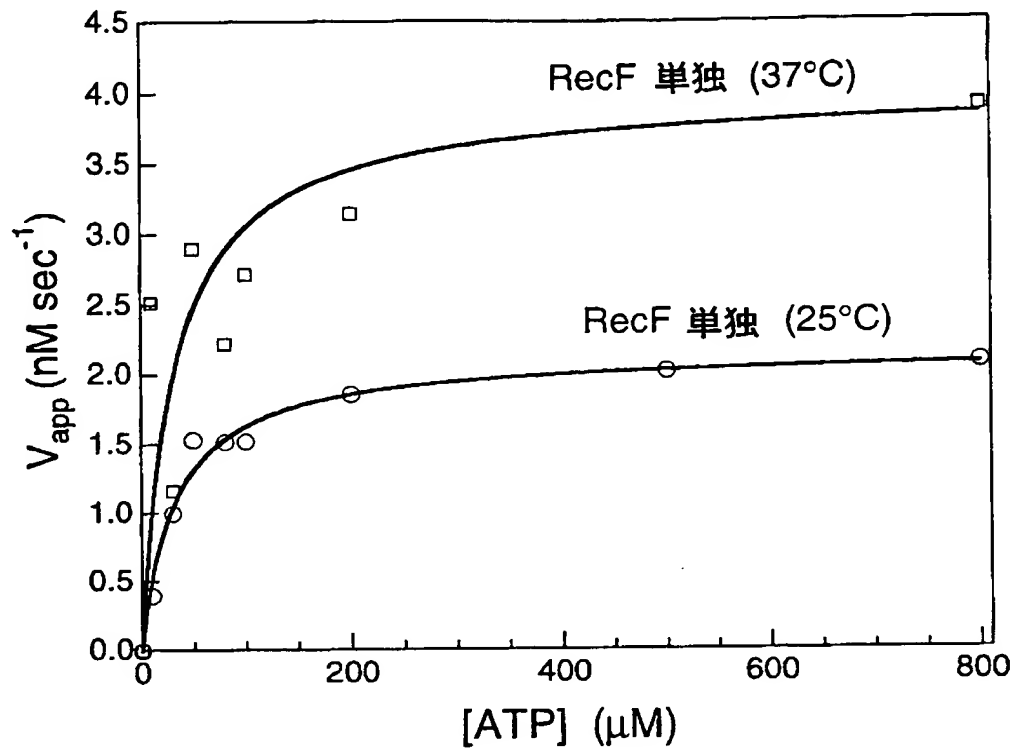
Tth	19	20	30	40	50	60
Eco						
Ppu						
Bsu						
Mtu						
Dra	MOD	VR	SA	L	TL	NR
Tth	70	80	90	100	110	120
Eco						
Ppu						
Bsu						
Mtu						
Dra	EQ	LI	R	A	V	RA
Tth	130	140	150	160	170	180
Eco						
Ppu						
Bsu						
Mtu						
Dra	W	I	P	E	D	L
Tth	190	200	210	220	230	240
Eco						
Ppu						
Bsu						
Mtu						
Dra						
Tth	250	260	270	280	290	300
Eco						
Ppu						
Bsu						
Mtu						
Dra						
Tth	310	320	330	340	350	360
Eco						
Ppu						
Bsu						
Mtu						
Dra						
Tth	370	380	390	343		
Eco						
Ppu						
Bsu						
Mtu						
Dra						

Tth: *Thermus thermophilus* HB8Eco: *Escherichia coli*Ppu: *Pseudomonas putida*Bsu: *Bacillus subtilis*Mtu: *Mycobacterium tuberculosis*Dra: *Deinococcus radiodurans*

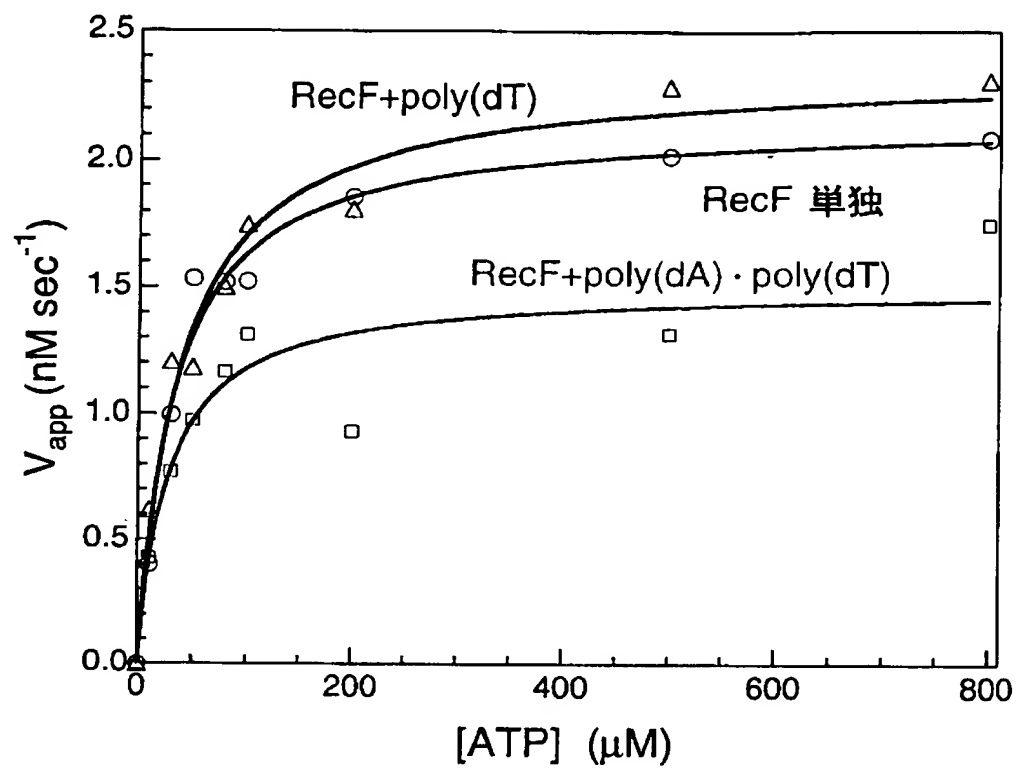
【図 28】



【図 29】



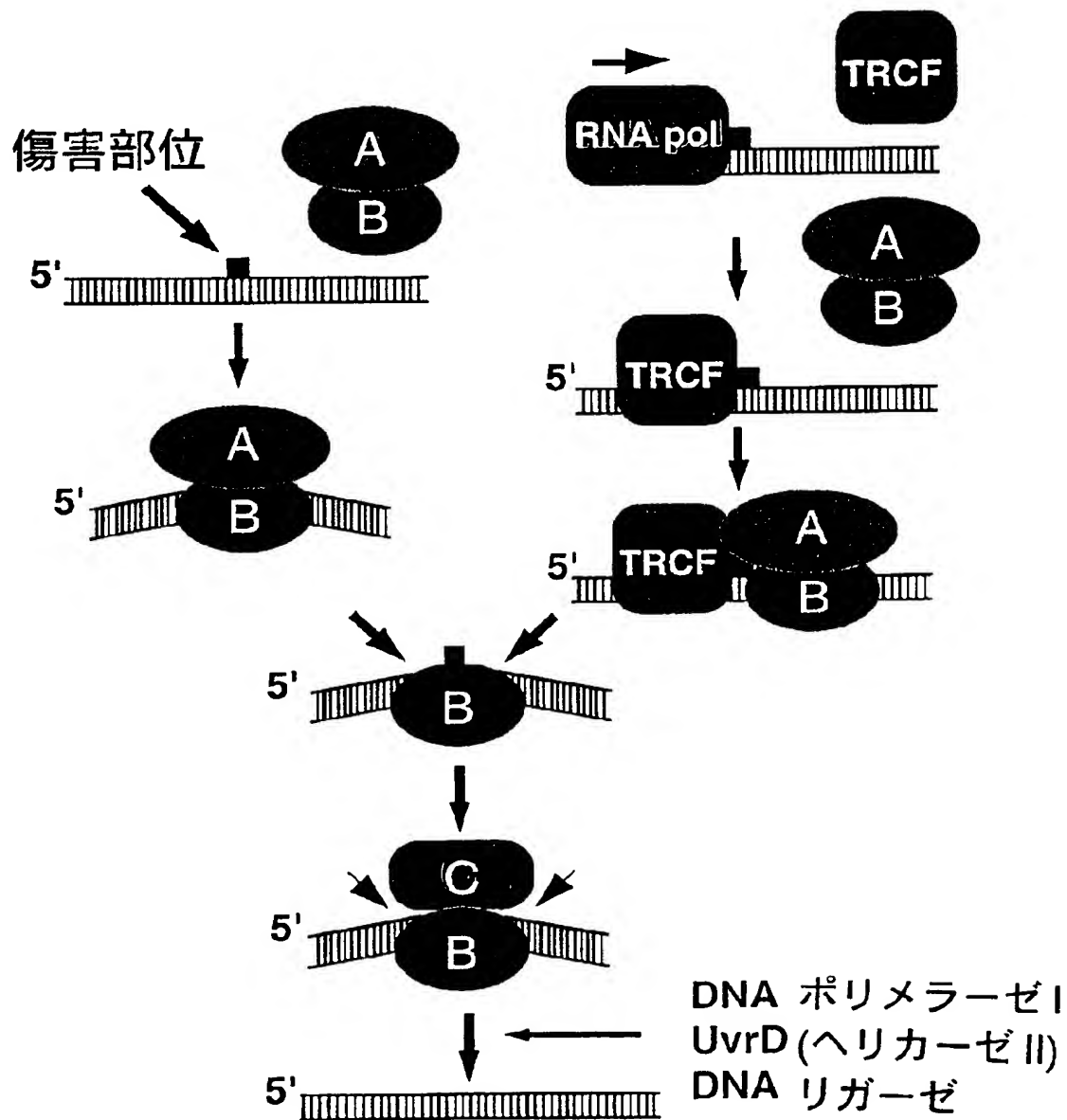
【図 30】



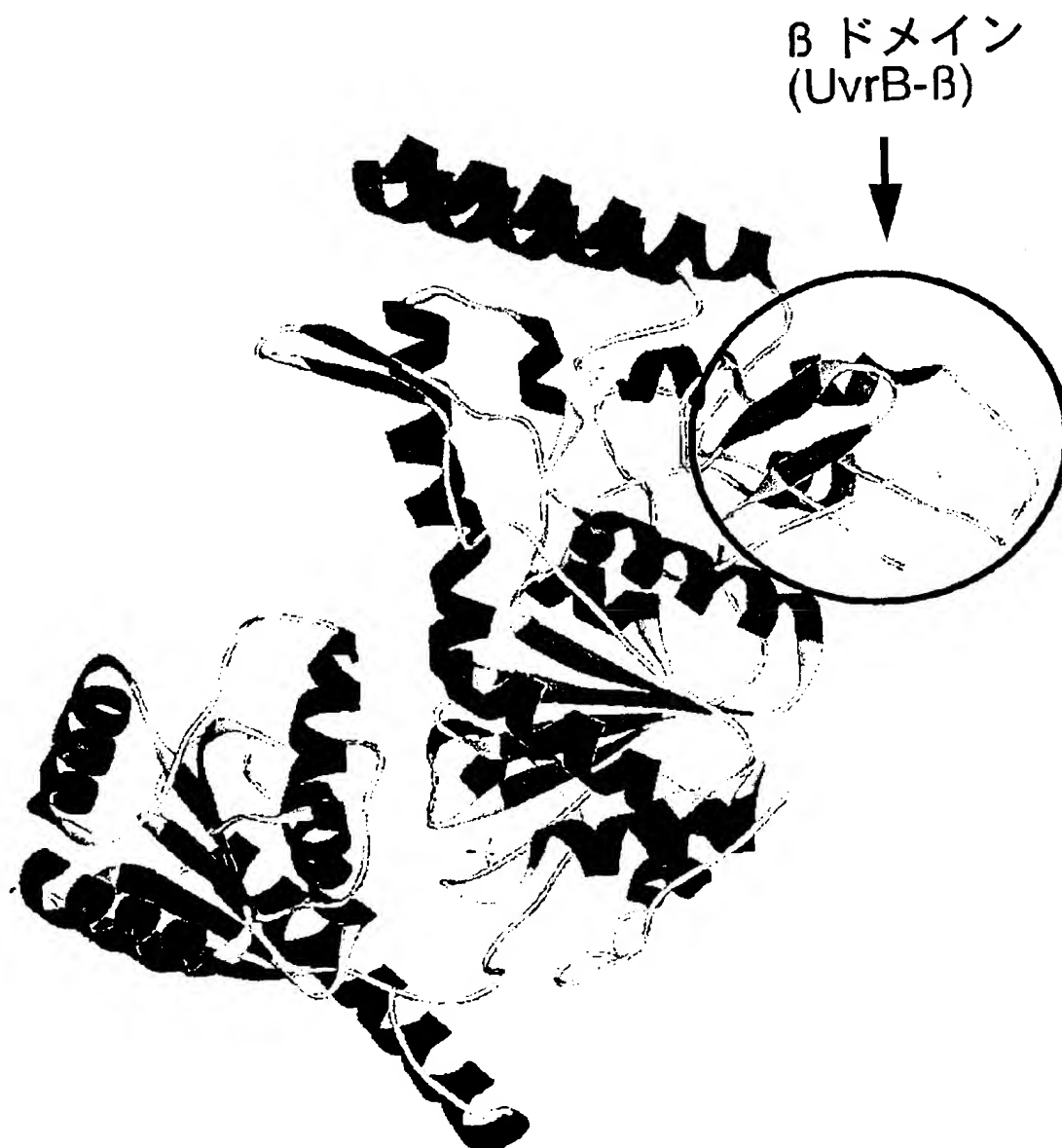
【図 3 1】

ゲノム全体の修復

転写共役修復

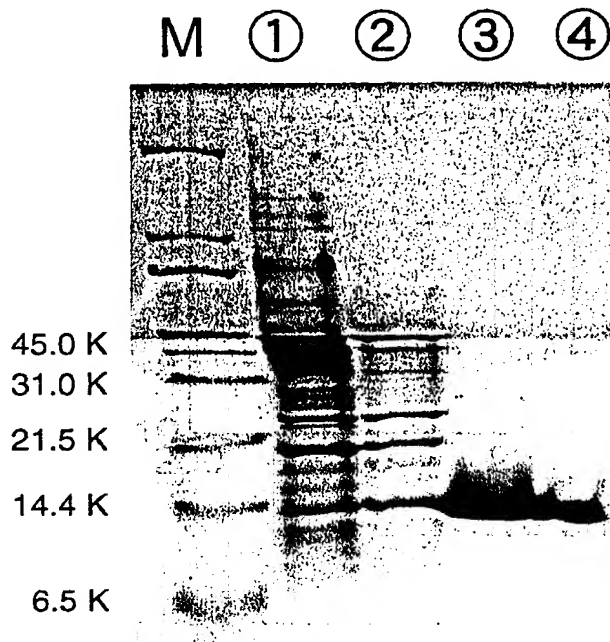


【図 32】

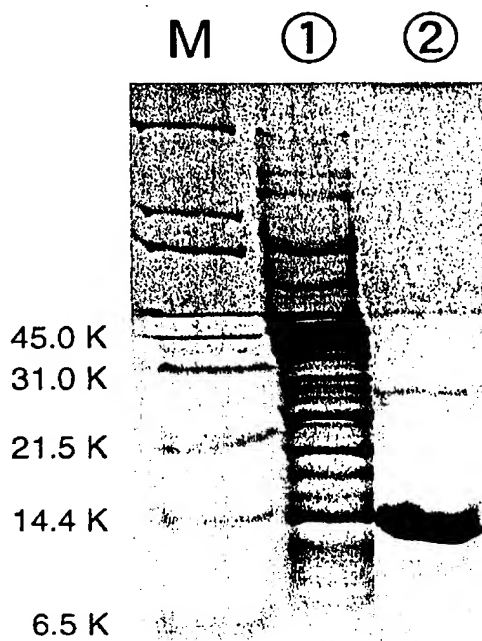


【図 33】

UvrB- β



TRCF- β



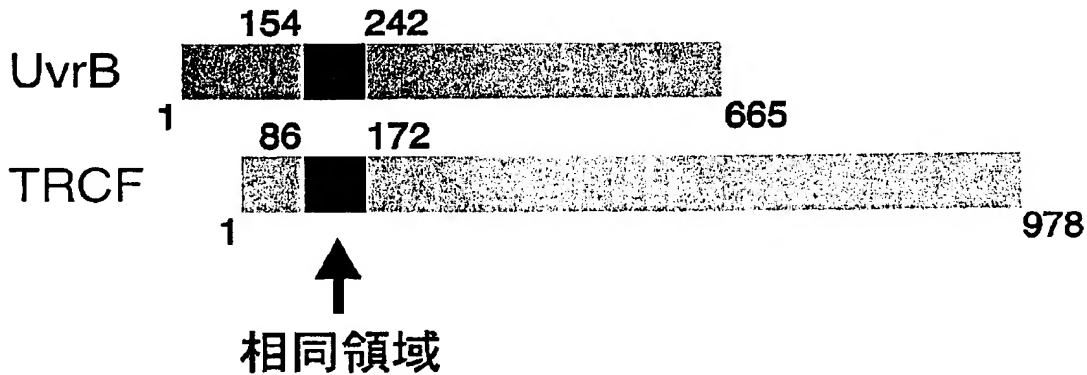
【図34】

UvrB-β 154 RNLVVERGKPYPREVLLERLLELGYQRNDI 184
 TRCF-β 86 WRLLLEVGRAYPREALLSRLLKLGYAR--- 113
 * . . * * . * * * * * * * * *

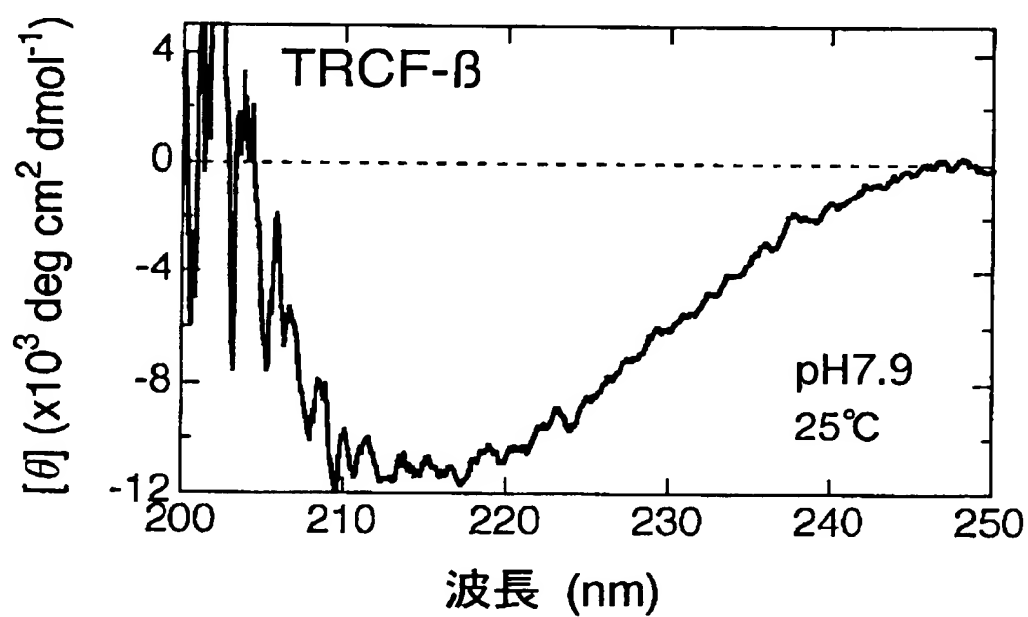
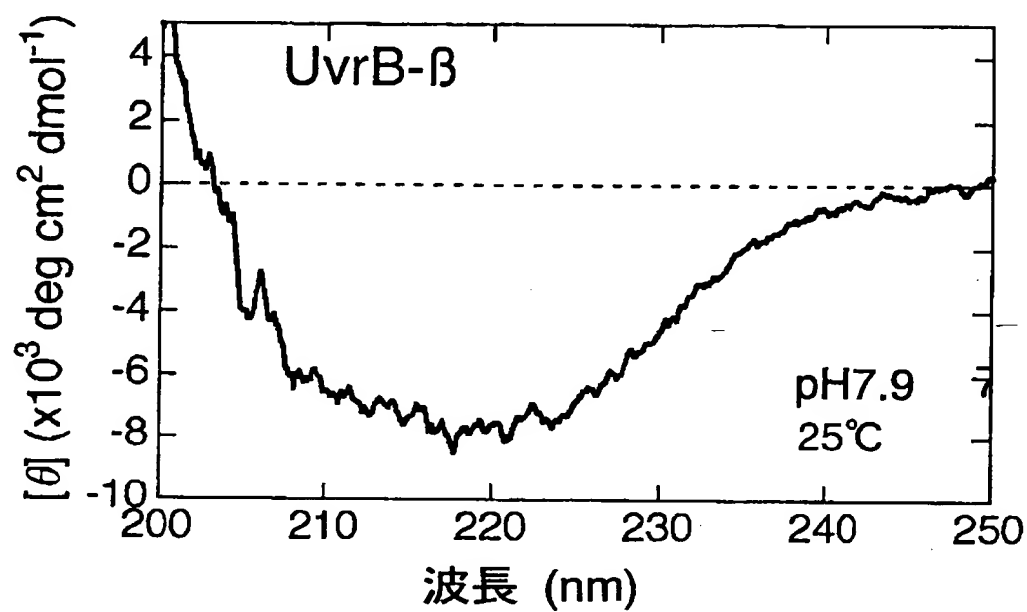
UvrB-β 185 DLSPGRFRAKGEVLEIFPAYETEPTRVELF 215
 TRCF-β 114 DED---YRVLGEVVELG-----EVRLEFF 148
 * . * * * * . * . * *

UvrB-β 216 GDEVERISQVHPVTG-ERLRELPG----- 236
 TRCF-β 149 GDELERLVVRGEERRRHVLLPKPGKAEGFT 163
 * * * * * * * *

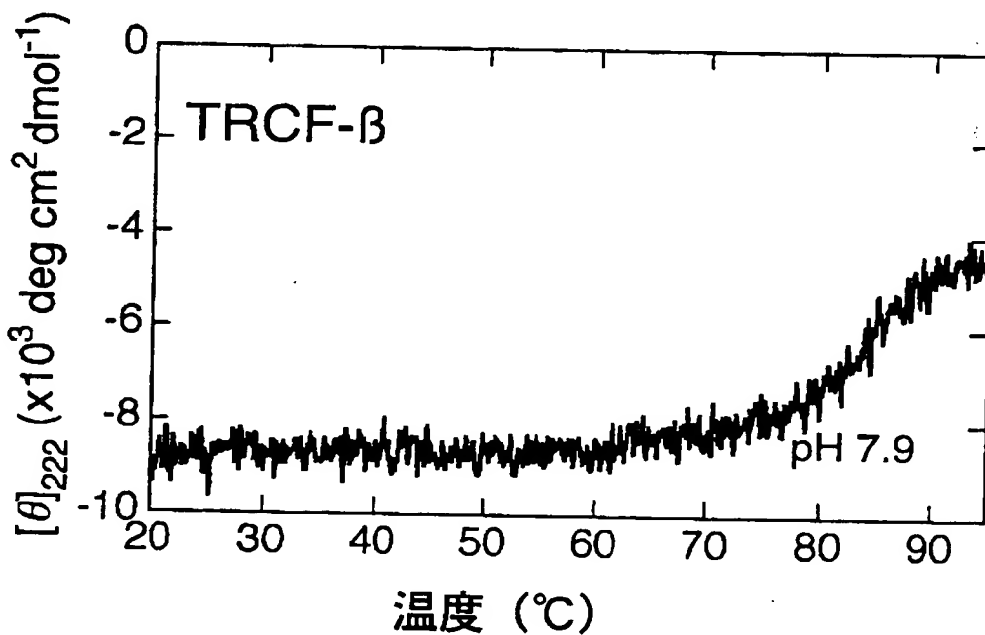
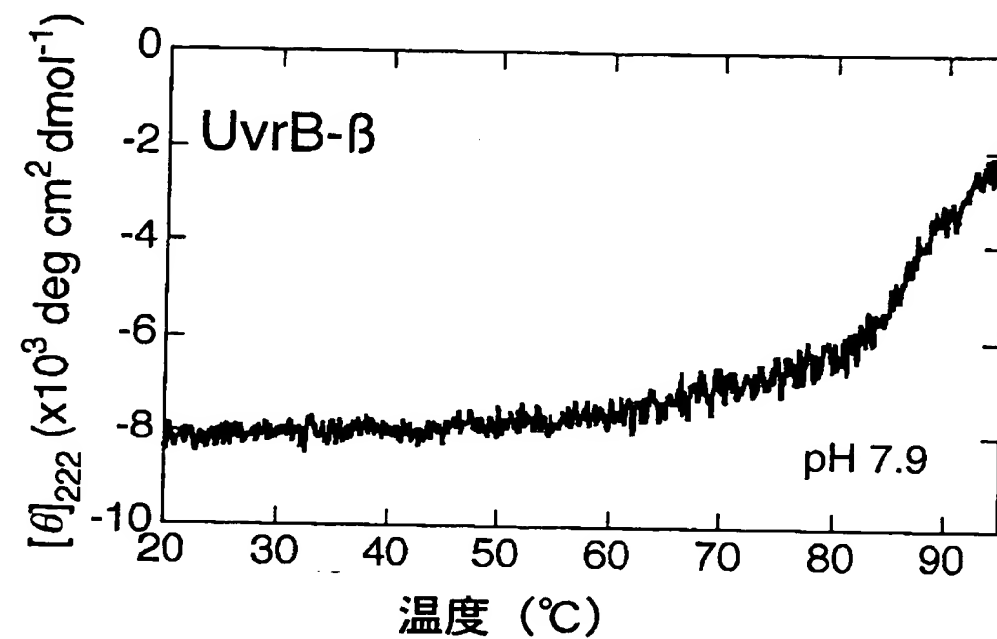
UvrB-β 237 ---FVL FPA 242 * 同一のアミノ酸残基
 TRCF-β 164 SKKVLHEPG 172 . 相同なアミノ酸残基
 . * *



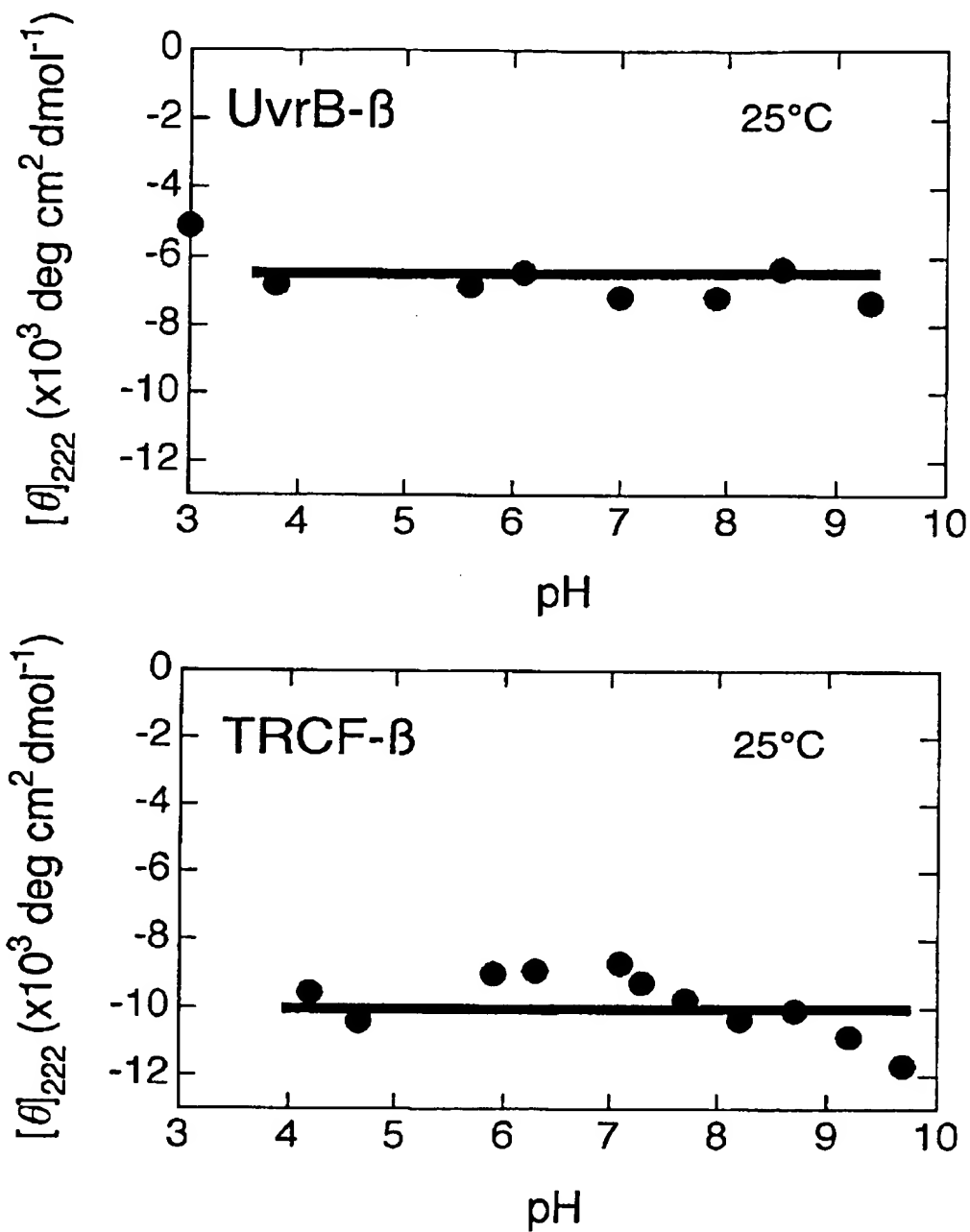
【図 35】



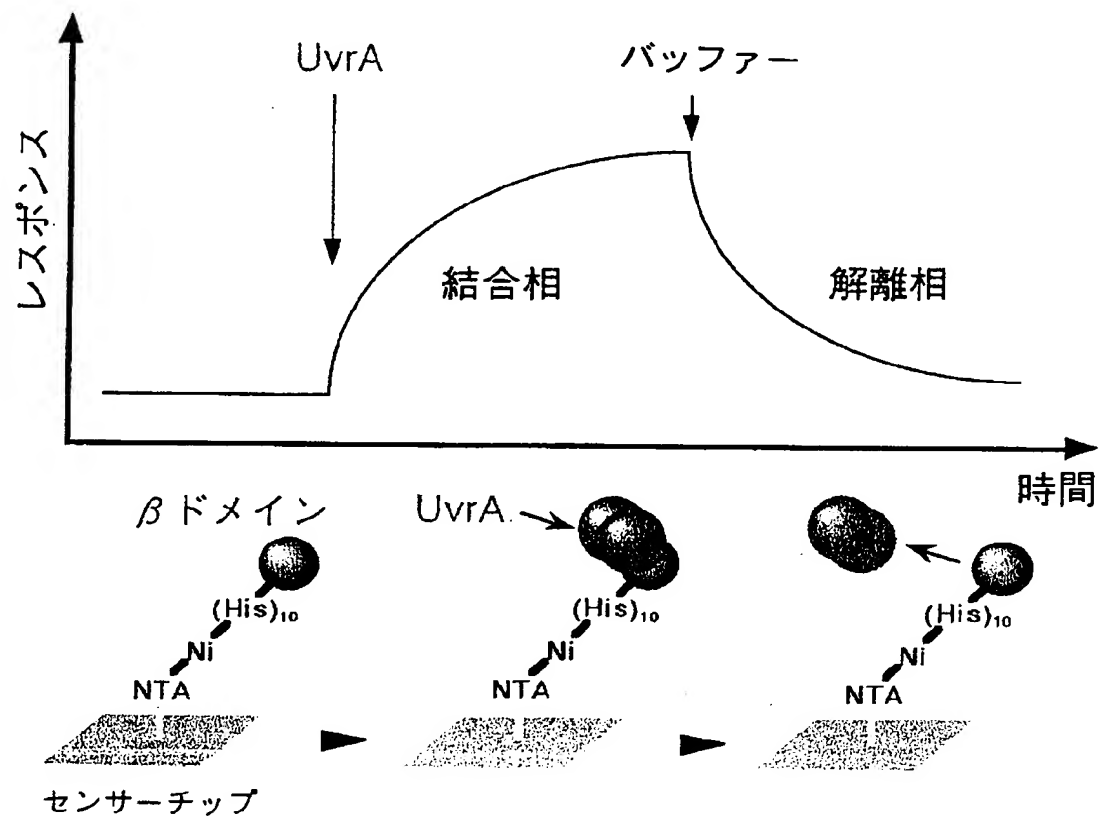
【図36】



【図 3 7】

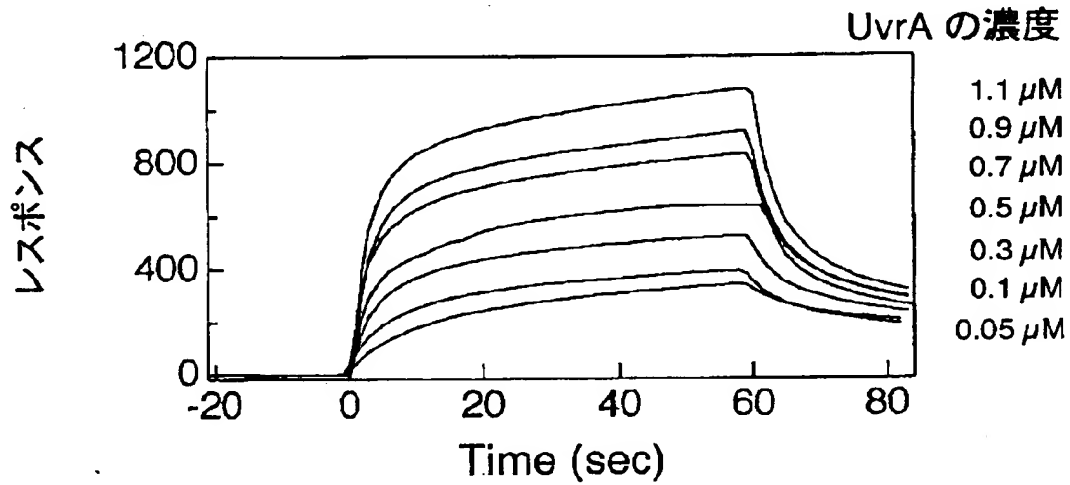


【図38】



【図 39】

センサーグラム



解析結果

	K_d ($\times 10^{-6}$ M)		k_{on} ($\times 10^5$ M $^{-1}$ S $^{-1}$)		k_{off} ($\times 10^{-1}$ S $^{-1}$)	
	- ATP	+ ATP	- ATP	+ ATP	- ATP	+ ATP
UvrB- β	2.6	0.4	2.0	1.5	5.2	0.6
TRCF- β	1.3	0.5	1.0	1.5	1.3	0.7

【書類名】 要約書

【要約】

【課題】 DNA修復酵素遺伝子の提供。

【解決手段】 以下の(a)又は(b)のタンパク質又は該タンパク質をコードするDNA修復酵素遺伝子。

(a) 配列番号 2、4、6 若しくは 8 に示されるアミノ酸配列を含むタンパク質

(b) 配列番号 2、4、6 若しくは 8 に示されるアミノ酸配列において1若しくは数個のアミノ酸が欠失、置換若しくは付加されたアミノ酸配列を含み、かつ、DNA修復酵素活性を有するタンパク質

【選択図】 なし

BEST AVAILABLE COPY

出 願 人 履 歴 情 報

識別番号 [000006792]

1. 変更年月日 1990年 8月28日
[変更理由] 新規登録
住 所 埼玉県和光市広沢2番1号
氏 名 理化学研究所